ARI Research Note 96-61

Novice Strategies for Comprehending Technical Texts

Diana Dee-Lucas and Jill H. Larkin
Carnegie Mellon University

Research and Advanced Concepts Office
Michael Drillings, Acting Director

April 1996

DTIC QUALITY INSPECTED 2

19960829 106

ARMY RESEARCH INSTITUTE

United States Army
Research Institute for the Behavioral and Social Sciences

DISCLAIMER NOTICE

UNCLASSIFIED

Technical Report
distributed by

DEFENSE
TECHNICAL
INFORMATION
CENTER

DTIC

UNCLASSIFIED

THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE
COPY FURNISHED TO DTIC
CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO
NOT REPRODUCE LEGIBLY.

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Carnegie Mellon University

Technical review by

Joseph Psotka

Novice Strategies for Comprehending Technical Texts

Table of Contents

Novice Strategies for Comprehending Technical Texts

Executive Summary

18 August 1989

Diana Dee-Lucas & Jill H. Larkin
Carnegie Mellon University

This project investigated the comprehension of technical texts by novice readers (i.e., people who are not familiar with technical subject matter). It focused on two questions: (1) how do novice readers determine what is important in technical texts, and (2) how does the organization of information in technical domains influence novice text processing and learning.

Our research on novice strategies for assessing importance found that novices develop rules defining what categories of information (i.e., definitions, facts, equations, etc.) are important in technical domains, and judge importance in technical texts using these rules. These rules determine what novices attend to during reading, what they remember later, and their depth of understanding of the text content. However, these rules are too general to allow novices to accurately assess importance in these texts. Consequently, novices miss some important content and devote effort to learning some of the less important information in these texts. Accordingly, these findings have practical implications for how technical texts should be written in order to correct for these importance rules and guide novices' attention to the appropriate text content.

Our research on the organization of information in technical domains showed that the order of mention of information in a text influences both attention and learning. When information is presented early in a text, it is more thoroughly processed and more likely to be recalled. This is because (1) readers expect important content to be presented first, and (2) there is greater uncertainty about the role of that information in the context of the passage as a whole. However, readers find some organizations more natural than others and will reorganize text content according to these preferred organizations. Additionally, novices find it easier to process technical texts if they present at the beginning general superordinate content that can be used as a organizational framework for processing later text information. These results have implications for how technical content should be organized for novice readers.

Novice Strategies for Comprehending Technical Texts

Final Report

Diana Dee-Lucas & Jill H. Larkin
Carnegie Mellon University

Research Problem

In technical domains, written text is the dominant mode of instruction, yet little is known about how novices process technical texts. Technical texts have unique characteristics that are unfamiliar to novice readers (i.e., readers who are not knowledgeable about technical subject matter), and therefore could pose comprehension problems. We have focused on two of these characteristics -- (a) complex content and (b) unique content organizations. The content of technical texts is complex and foreign in nature to novices, and novices may therefore have difficulty picking out the important information in these texts. Therefore, the first focus of this research was on how novice readers determine what is important in technical texts. Additionally, technical texts use specialized organizational schemes in structuring their content, and novices may have difficulties detecting and using these novel organizations while they are reading. The second focus of this research was on how one such specialized text structure influences novice attention and learning. This project has determined how both of these unique features of technical writing influence novice text processing, and has developed corresponding suggestions for the design of technical texts.

Overview of Major Findings

Novice strategies for assessing importance. This research found that novice readers judge importance in technical texts partly on the basis of information category -- they consistently judge some categories of information, such as definitions and equations, as being more important than other categories, such as facts. This indicates that novices develop general rules defining what types of information are important in technical domains. Further research showed that these rules have a significant impact on learning, in that they influence what novices attend to, what they recall, and ultimately their depth of understanding (as indicated by their ability to use information to solve problems). Given the influence of these rules on learning, the accuracy of these rules is of great importance in determining the efficiency with which novices learn from technical texts. Previous research comparing experts' and novices' judgements of what is important in technical texts indicates that these rules are too general to allow novices to accurately assess importance in these texts. Accordingly, these findings have practical implications for how technical texts aimed at a novice audience should be written in order to compensate for incorrect novice rules and capitalize on those that are appropriate.

Text organization and learning. This research compared novice processing of a specialized text structure with a more conventional alternative structure. The specialized structure was one used in presenting principles and their corresponding proofs. This structure presents the proof prior to stating the principle, a "proof-first" organization. This was compared to the more familiar "principle-first" organization in which the principle is presented prior to the proof. The results indicated that novices find .the specialized proof-first structure less appropriate and more difficult to process. Novices used the conventional principle-first structure in summarizing both proof-first and principle-first texts, indicating that they clearly preferred this structure. Additionally, the proof-first structure decreased the probability of readers recalling the passage theme (i.e., the principle). Finally, the two structures caused novices to focus attention on different types of information, with novices giving the most attention to information when it occurred at the beginning of the text. These results suggest that the specialized proof-first structure makes it more difficult for novices to develop an overall representation for the text, and ultimately results in a more fragmented representation which decreases recall of the principle. These results have implications for how technical content should be organized for novice readers.

Potential Applications

Much of the Army's instruction of novice personnel on the use of complex equipment occurs via technical manuals. The findings from this research project have two main implications for the improvement of technical manuals.

First, our research found that novices with little technical background have strong preconceptions about the types of information that are important in these texts. These preconceptions are misleading in that they result in novices focusing on some of the less important text content. This suggests that the effectiveness of technical manuals can be improved by taking these preconceptions into account and using clear signals to indicate important material that novices might miss. This can be done in part by manipulating the form in which information is presented. Because novices have rules about what categories of information are important, they are sensitive to how information is presented (i.e., whether information is presented in the form of a definition or a fact). Therefore, writers can use sentence form to guide novices to the important content in texts. For example, important quantitative relations should be presented as equations because novices consider equations to be a particularly important type of information. Similarly, unimportant quantitative relations (i.e., the intermediate equations in a proof, or low-level elaborative definitions) should be presented in non-equational form, or otherwise de-emphasized. The latter can be done, for example, by not setting the equation apart from the rest of the text (presenting it on the same line with text), or by not numbering less important equations. Additionally, writers should take care to emphasize important content that novices are likely to miss because they consider that content to be a less important type of information (for example, facts). This information should be presented in a manner that calls attention to it, either by signally its importance through typographic cues (e.g., italicizing, underlining, etc.), or by explicitly indicating indicating its importance in the text (e.g., "you should note that," "it is important to understand," etc.). In general, writers can use knowledge of novices rules in conjunction with rhetorical indicators of importance to emphasize important content, de-emphasize less critical information, and thus guide attention to the appropriate text content.

Second, our findings show that the processing difficulty of technical texts depends in part on how the content is organized. Processing difficulty is reduced when major principles are presented early in the text. Apparently this type of content provides novices with an organizational framework which aids them in organizing the more detailed text information and helps them determine what is important as they read. This research also indicates that writers need to match the text organization that they use to their learning goals for the reader. The information that is most important for the reader to learn will receive the most attention if it is presented early in the text. However, if the text structure does not match an organization that is natural for the reader, the reader may tend to reorganize it by placing the information they consider to be most important at the beginning of the text.

Summary of Research

Strategies for assessing importance in technical writing. Technical texts are typically densely packed with complex information, including equations, symbols, and specialized terms. Consequently, it can be difficult for readers who are unfamiliar with technical subject matter to pick out the important information. We conducted a series of experiments examining the basis on which novice readers determine what is important in technical texts. We found that novice readers judge importance partly on the basis of information type (i.e., definitions, facts, equations, etc.). They develop rules defining what types of information are important in technical domains. We demonstrated this in several experiments in which we had novices and experts judge the importance of information when it was presented in different forms (see Appendix B -- Dee-Lucas & Larkin, 1988b). In one experiment, subjects judged the importance of content presented either as a definition or a fact. In another, subjects judged the importance of quantitative relations presented either as equations or written out in sentence form. The content was identical in both versions--only the form of the content varied (i.e., definition vs. fact, or equation vs. sentence).

In both studies, novices were influenced by sentence form in judging importance. In the first experiment, novices considered information to be more important when it was presented as a definition rather than a fact. In the second study, novices thought the content was more important when it was presented as an equation as opposed to a written statement. Experts, on the other hand, were not influenced by sentence form in judging importance. They judged the same content as equal in importance regardless of how it was presented. These results show that novices develop rules defining what general categories of information are important in technical domains, and they judge importance according to the category-membership of the content rather than on the basis of the content itself.

Based on these initial findings, we conducted follow-up research to determine how these category rules influence novice text processing and learning. This research indicated that these rules influence attention and memory for technical texts, as well as problem-solving ability. Novice readers spend more time on the types of content that they consider to be important when studying technical texts, and also recall more of this content (see Appendix A -- Dee-Lucas & Larkin, 1988a). As a result, these rules ultimately influence novices' ability to use technical content in problem solving (see Appendix D -- Dee-Lucas & Larkin, 1989b). This was demonstrated in an experiment in which we compared novice comprehension of texts presenting technical proofs when the proofs were presented as a series of equations (a type of content that novices consider to be particularly important) as opposed to a series of written statements equivalent in content to the equations. We found that with the equation-based proofs, novices tend to focus on quantitative content to the exclusion of other information types, so they miss important information linking together the equations into a coherent proof. Consequently, they have difficulty solving problems requiring knowledge of non-quantitative proof content. Additionally, novices focus on equations for the purpose of memorizing them, without really trying to understand them. As a result, novices recall equations well, but have only a superficial understanding of their meaning so that novices are unable to correctly apply equations in solving problems.

These findings indicate that novice importance rules have a major impact on novice text learning. Thus the accuracy of these rules is important in determining the efficiency with which novices learn from technical texts. Previous research comparing novices' and experts' judgements of what is important in technical texts indicates that these rules are too general to allow novices to accurately determine what is important. Therefore, the findings from the current research indicate that texts aimed at a novice audience must take into account novice importance rules and include clear indicators of importance that aid novices in distinguishing the important from the less important content.

<u>Text organization and novice comprehension.</u> Different content domains have different organizations that are used to structure information in that domain. These include, for example, the structure of legal documents and scientific reports. These specialized structures are specific to the particular domain, and hence unfamiliar to novice readers. Technical texts also make use of specialized organizations in presenting technical content. We examined the effects of one of these specialized structures on novice comprehension of scientific proofs and principles. Technical texts often introduce a new principle by presenting its proof prior to stating the principle, so the reader has no idea of the point of the proof at the time it is being read. This type of structure is unfamiliar to novices, and lacks a conceptual framework that readers can use in processing the text. Thus novices might have difficulty learning from texts having this kind of structure. We conducted a series of studies in which we compared novices' processing of texts having this specialized "proof-first" structure with texts having the more conventional "principle-first" organization (i.e., the principle presented prior to its proof). The principle-first structure is both more familiar and provides readers with thematic information (i.e., the principle) to guide their processing of the text. We conducted three studies: one looked at differences produced by these two structures in reading times and summaries, another looked at their influence on the perceived importance of the information, and the third examined differences in what was recalled. We found that these two structures produce differences in how the content is processed, organized, and recalled.

Readers spend more time on text information when it is presented first. Thus more time is spent on proof information when it is in the proof-first format, and more time is spent on the principle with the principle-first organization. The importance rating study showed that there are two factors contributing to this reading time effect. First, readers consider content to be more important when it is presented first as opposed to being placed later in the text. Thus readers spend more time on this content because they think it is particularly important to learn. Second, although they consider this information to be particularly important, readers are less certain about the precise importance of the content when it is presented first. Readers have more difficulty predicting the exact importance of information (on a scale of 1 to 5) when the content occurs early in the text. Thus readers spend more time on beginning content because they are less certain of its precise importance in the text.

Although the two structures focused attention on different types of information within the texts, readers found the proof-first structure more difficult to process overall. They had more difficulty predicting the importance of information with the proof-first structure, suggesting that readers experienced more confusion when reading with this structure. With the principle-first structure,

4

readers were apparently able to use the principle as an organizational theme for guiding their processing of the text as they were reading.

Text structure influenced memory for the principle, but not the proof. Readers were more likely to recall the principle both immediately and after a delay with the principle-first structure; recall of the proof did not vary with structure. Thus the proof was equally memorable with both structures, even though its processing was facilitated by the principle-first structure. There are several possible reasons why the principle-first structure would enhance recall of the principle. First, this structure allows readers to use the principle as a conceptual framework for processing the text. Therefore, readers keep reviewing the principle mentally as they read in order to interpret the rest of the text in terms of its relationship to the principle. Second, the principle was repeated in the principle-first texts in a manner the allowed them to review this content in an optimal manner (i.e., spaced apart in the text). With the proof-first structure, the review of the principle was grouped in one portion of a text, a format that has not been shown to be optimal in previous research.

The readers clearly prefer the principle-first structure, in that they use this structure in summarizing both proof-first and principle-first texts. In summarizing the proof-first passages, readers reorganize the proof-first texts into a principle-first structure. The summary data suggests that readers may re-structure texts according to an organization that they feel is most appropriate given what they consider to be the most important text content.

The results of this research suggest that the structure used by a writer should match the instructional goals of the text. If the main goal is to have the learner understand the principle, and the proof is provided as additional elaborative information designed to strengthen understanding of the principle, then the principle-first structure seems most appropriate. This structure results in the desired learning outcomes (as indicated by this research) and corresponds to novices' view of how this type of content should be organized, most likely because it is consistent with writing conventions regarding the placement of topic information. In this sense, it corresponds to an expository text schema that novices have for this type of content. Understanding the principle is the most common learning goal associated with this type of content. However, other goals are possible for which the proof-first structure might be more appropriate. One would be if the main topic of interest were the methodology underlying the proof; another would be if the logic behind the proof were of most concern. In these cases the nature of the proof would be of greater interest than the specific principle. Thus a structure analogous to an argument leading up to a conclusion would be most appropriate in that it would focus attention on this content. However, the current research indicates that in these cases, providing an informative introduction to the proof containing appropriate orienting information would help readers process texts having this type of structure.

Contributions to Basic Science

These results suggest how novices develop a "content schema" for a new knowledge domain. A content schema indicates how the knowledge for a particular domain is typically structured, including what is important in the domain. This research suggests that an early stage in the development of a content schema may be the specification of rules indicating what categories of information are important. When first learning about a subject matter, novices discover that certain easily identifiable types of information (e.g., equations and definitions) are likely to be important, and form general importance rules on this basis. These rules could form the foundation for developing an expert content schema by providing a relatively undifferentiated base from which a more refined rule structure could be developed. In developing an expert content schema, novices would move from an importance-rule system based on information categories to a more refined knowledge structure based on a deeper analysis of the nature of the content. Thus novice importance rules can be thought of as the rudimentary beginnings of a content schema for technical domains.

This research also indicates that the order of mention of information in a text influences both attention and learning. When information is presented early in a text, it is more thoroughly processed and more likely to be recalled. This is because (1) readers expect important content to be presented first, and (2) there is greater uncertainty about the role of that information in the context of the passage as a whole. However, readers find some organizations more natural than others and will use these organizations preferentially over those provided for them. This indicates that readers not only use text structure as an indicator of importance, but also have preconceptions about what the critical content should be, which was also shown in the research on novice importance rules.

Reports Resulting from Contract

Dee-Lucas, D. (1985). *Attentional strategies for studying scientific texts.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Dee-Lucas, D. (1986). *Novice rules for assessing importance in scientific texts.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Dee-Lucas, D. & Larkin, J. H. (1986). *Novice rules for assessing importance in scientific texts.* (Tech. Rep. A.C.P. 16). Carnegie Mellon University.

Dee-Lucas, D. & Larkin, J.H. (1987). *Attentional strategies for studying scientific texts.* (Tech. Rep. A.C.P. 17). Carnegie Mellon University.

Dee-Lucas, D. (1987). *Sentence form effects in novice importance judgements.* Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Dee-Lucas, D. & Larkin, J.H. (1987). *Novice importance rules: Definitions and equations.* (Tech. Rep. A.C.P. 18). Carnegie Mellon University.

Dee-Lucas, D. (1988). *Organization and comprehensibility in scientific texts.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Dee-Lucas, D. & Larkin, J.H. (1989). *Organization and comprehensibility in scientific proofs, or "Consider a particle p...".* (Tech. Rep. A.C.P. 19). Carnegie Mellon University.

Dee-Lucas, D. & Larkin, J.H. (1989). *Equations in scientific proofs: Effects on comprehension.* (Tech. Rep. A.C.P. 20). Carnegie Mellon University.

Publications Resulting from Contract

Dee-Lucas, D. & Larkin, J.H. (1988a). Attentional strategies for studying scientific texts. *Memory & Cognition,* **16**, 469-479.

Dee-Lucas, D. & Larkin, J.H. (1988b). Novice rules for assessing importance in scientific texts. *Journal of Memory and Language,* **27**, 288-308.

Appendices (Major Reports & Publications Resulting from Contract)

Appendix A: Dee-Lucas, D. & Larkin, J.H. (1988a). Attentional strategies for studying scientific texts. *Memory & Cognition,* **16**, 469-479.

Appendix B: Dee-Lucas, D. & Larkin, J.H. (1988b). Novice rules for assessing importance in scientific texts. *Journal of Memory and Language,* **27**, 288-308.

Appendix C: Dee-Lucas, D. & Larkin, J.H. (1989a). *Organization and comprehensibility in scientific proofs, or "Consider a particle p...".* (Tech. Rep. A.C.P. 19). Carnegie Mellon University.

Appendix D: Dee-Lucas, D. & Larkin, J.H. (1989b). *Equations in scientific proofs: Effects on comprehension.* (Tech. Rep. A.C.P. 20). Carnegie Mellon University.

Appendix A

(Total pages in length: 11)

Dee-Lucas, D. & Larkin, J.H. (1988a)

Attentional strategies for studying scientific texts.

Memory & Cognition, **16**, 469-479

Attentional strategies for studying scientific texts

DIANA DEE-LUCAS and JILL H. LARKIN
Carnegie-Mellon University, Pittsburgh, Pennsylvania

Content-area novices develop rules as to what types of information (e.g., definitions, facts, equations) are important in texts (Dee-Lucas & Larkin, 1988). The study reported here indicates that these rules influence text learning. Experts and novices read and recalled science texts. Reading times and recall for different types of content were compared for the two groups. Results indicate that novices' importance rules function as part of novices' control schema during reading, influencing their attentional processes and the resulting representation formed for the text. This was evident in qualitative differences between experts and novices in their attentional patterns and text recall. The study also found that the number of passage readings and the passage topic have a greater influence on the reading times of experts, but both groups adjust processing time according to the hierarchical level of the passage content. The findings are discussed in terms of their implications for novices learning from texts.

Recent research on text comprehension has explored how readers' knowledge structures interact with text content in determining what is learned. These knowledge structures range from very general structural schemata for a writing genre (e.g., story grammars) to domain-specific representations of a content area (e.g., the goal structure for baseball). They all aid readers' comprehension by guiding attention to important and relevant content, facilitating encoding of that information into existing knowledge structures, and providing retrieval cues. Thus a reader's prior knowledge plays an important role in determining what is learned from text.

The present study examined how the primitive knowledge structures of novices influence attention and learning from unfamiliar scientific texts. Most research on knowledge effects has considered well-developed stable knowledge structures, such as story schema, or expert knowledge. Relatively little is known about the more tentative knowledge structures developed by novices as they learn about a content domain. These knowledge structures are difficult to characterize due to their transient nature (i.e., they are continually updated and modified as more knowledge is acquired) and the potential variability among people beginning to learn about a content area. However, Dee-Lucas and Larkin (1986, 1988) have begun to describe the knowledge structures of novices learning physics. They have found that people learning physics form rules that indicate what types of easily recognizable information (e.g., definitions, facts, equations) are important in physics texts. These rules are part of novices' developing knowledge structure for the domain of physics.

The current study investigated the extent to which these importance rules influence novices' attention and learning when they are studying physics texts (i.e., repeatedly rereading passages to prepare for a test on content). The research examined whether novice rules are used to guide attention during reading, how the use of these rules interacts with overall passage familiarity in modifying attention during repeated study trials, and whether the resulting attentional differences are reflected in what is remembered afterward. To the extent that novice importance rules control text processing, the nature of these rules (i.e., their initial accuracy and how they change with additional knowledge) has important implications for the efficiency with which novices learn from texts, and how this efficiency can be increased.

Although recent research has emphasized the importance of domain-related schematic knowledge in text comprehension, readers seem to be able to identify important content in texts about unfamiliar topics without using this type of knowledge. Kieras (1980, 1985; Kieras & Bovair, 1981) showed that readers can identify thematic information in texts by using surface-level text features (e.g., the topic-comment structure of sentences, passage organization, signaling phrases) and a limited, or "shallow," understanding of the semantic relations among the content (e.g., using semantically entailed information from familiar terms to infer the nature of unfamiliar terms). Thus readers seem to be able to identify the main points in unfamiliar technical texts without relying on knowledge about the content domain by using superficial characteristics of the text structure and semantics. Kieras (1985) argued that domain-specific schemata, therefore, are not particularly important for identifying the important content in technical prose.

However, Dee-Lucas and Larkin (1986, 1988) showed that domain-specific rules are used by novices in assessing importance in texts. They found that content-area novices (i.e., beginning-level students) develop rules

about the relative importance of different categories of information in a domain. These rules are domain specific in that naive subjects (i.e., subjects without domain training) do not possess these rules. Dee-Lucas and Larkin (1988) demonstrated this rule use in a series of experiments in which they manipulated the category membership of information without altering its substantive content and had subjects varying in expertise judge its importance. They found that naive subjects judged the same content as equal in importance regardless of its sentence-category membership. Similarly, experts also judged importance independently of category membership. However, novices altered their assessment of the importance of the passage content according to its category. For example, they judged the same content as more important when it was presented as a definition than as a fact. Thus, although naive, expert, and novice subjects are all able to identify definitions and facts, category membership is only relevant to the novices in assessing importance; that is, only novices think that one category type is more important than another. This is because novices have developed rules regarding the importance of the different categories of information as a result of their limited instruction in the content domain.

These novice importance rules can be thought of as the rudimentary beginnings of a content schema. A content schema indicates how information in a content area is typically organized, including what is important in that domain. These rules are part of novices' evolving content schema, in that they define in part what novices view as important. Sentence category is irrelevant for people without domain training because they lack a content schema and do not have strong expectations about what types of information are important. Sentence category is also irrelevant for experts, who have rich content schemas that allow them to judge importance directly. An evolving novice schema would consist of only a few category rules that specify the importance of different information types. However, with additional knowledge regarding the relative importance of information within type categories, these rules could be refined and develop into a differentiated expert knowledge system. Because novice knowledge structures are primitive, with only a few general rules, novices systematically misidentify the important content in domain-related texts. Comparisons of experts' and novices' judgments of what is important to learn in natural texts indicate that novices do not discriminate among the important and less important content within type categories to the same degree as do experts (Dee-Lucas & Larkin, 1986).

The current study was an attempt to determine whether the primitive content schema of novices (i.e., sentence-type rules for judging importance) is used by novices in determining what is important as they read, as indicated by the amount of processing time given to different types of information. Reading times for different types of content were examined to see if novice rules are reflected in differential attention given to different category types.

Additionally, the amount of information recalled from different category types was examined to see if level of recall varies with the judged importance of the category, that is, if attentional differences produced by importance rules also are reflected in qualitative differences in recall.

Kieras's (1980, 1985; Kieras & Bovair, 1981) findings, along with the results of research on novice importance rules, indicate that novices have available several sources of information for judging importance: domain-independent properties of the text structure and content, and primitive domain-specific schemata arising from limited experience with the text domain. Given that novices have available alternative sources of information for judging importance, they might vary the degree to which they rely on these sources when studying a text (i.e., repeatedly rereading a passage to learn its content). Novices might be more likely to rely primarily on objective text-based cues on the initial reading of an unfamiliar passage, and to identify the main ideas using rhetorical indicators of importance. Once they were familiar with structural characteristics of the passage (i.e., its overall organization and emphasis), they could then begin to make more use of their schematic knowledge (i.e., importance rules) on subsequent passage readings and concentrate on parts of the text defined as important within this schema.

In the current study, we investigated this possibility by having subjects read each passage three times, examining reading times for each reading trial. The experimental situation was designed to be similar to a study situation. The subjects were told that they were to read each passage three times and that they would receive a test on its contents. Reading times for initial and subsequent readings were compared to determine whether importance rules interact with overall passage familiarity in controlling attention.

This study focused on attentional and recall differences for definitions and facts, two easily recognized types of information that are common and important in scientific texts. The results of earlier research using naturalistic texts indicate that both novices and experts judge definitions as more important to learn than facts (Dee-Lucas & Larkin, 1986). However, novices are even more likely than experts to judge definitions as important and facts as unimportant.[1] Thus novices are missing some important facts and judging as important some less important definitions. This indicates that novices do not distinguish between the important and unimportant content within these type categories. As noted earlier, novices (but not experts) judge definitions as more important than facts, even when the content of the two sentence types is held constant in experimental texts (Dee-Lucas & Larkin, 1988).

Because the definitions and facts in the passages used in the current study differed in content, they varied in syntactic complexity and familiarity. Content familiarity has been shown to influence both reading time (Graesser, Hoffman, & Clark, 1980; Johnson & Kieras, 1983) and recall (Chiesi, Spilich, & Voss, 1979; Johnson & Kieras, 1983; Spilich, Vesonder, Chiesi, & Voss, 1979).

Familiarity and syntactic complexity were controlled in the data analyses with individual subject ratings. These were performed after the subjects had finished reading each passage and completed a free recall test on its content. Each subject then rated the target sentences according to (1) how familiar he/she had been with the sentence content prior to reading the passages, and (2) how comprehensible the sentences were independent of the familiarity of the content. These ratings were used in the analysis of the reading times to control for differences due to these two variables.[2] The familiarity ratings were also used as a covariate in the analyses of the recall data.

The influence of passage structure on reading times for definitions and facts was also controlled in this study. Dee-Lucas and Larkin (1986) found that the difference in the judged importance of the two sentence types varied with the hierarchical level of the content. To determine whether importance rules and text structure interact in influencing attention, we analyzed the organizational structure of the passages in this study into a heirarchy, and classified definitions and facts according to their hierarchical level. The reading times were analyzed as a function of both information type and hierarchical level to determine if the difference between the time spent on definitions and facts varied with their structural importance. This analysis also provided an opportunity to examine general levels effects in reading time.

METHOD

Stimulus Materials

The two passages were those used by Dee-Lucas and Larkin (1986) in their research comparing expert and novice importance ratings. The topics were work and energy, and fluid statics. They were 44 and 47 sentences in length. Each passage had been constructed so that it contained 18 target sentences, 9 definitions and 9 facts. The definitions described a concept or construct in terms of a given set of characteristics. The facts presented properties of constructs or substances, such as "The amount of force produced by a fluid varies with surface area," "Atmospheric pressure also varies from day to day due to changes in the air temperature," and "The pressure exerted by the fluid on each part of the surface of this submerged body does not depend on the material that the body is made of."

The definitions and facts were classified according to their level in the passage structure. The structural analysis was performed at the sentence level rather than the more commonly used proposition level (Kintsch, 1974; Kintsch & van Dijk, 1978; Meyer, 1975). This is because the dependent measures used in this and prior research on novice importance rules have involved complete sentences. The structural analysis (also reported in Dee-Lucas & Larkin, 1986) produced a hierarchy, with the major points of the passages occurring at the top levels and the more detailed information modifying these points occurring at the lower levels. This analysis was performed by first selecting from the passages the main topics or concepts and placing them at the top of the hierarchy. All sentences whose primary content modified these top-level statements were placed at the next level; information modifying second-level content was placed at the third level. Second- and third-level sentences consisted of examples, derivations (i.e., information implied

by higher level content), explanations, subtopics, preconditions (i.e., conditions necessary for a rule, principle, or fact to hold true), attributes, and properties. The target sentences were six facts and six definitions occurring at each of the first three hierarchical levels of the passage hierarchies. Importance rating data reported in Dee-Lucas and Larkin (1986) indicate that the judged importance of the target sentences is related to their hierarchical levels, as specified by this analysis procedure. One of the passages is included as an Appendix, with the target definitions and facts noted, along with their hierarchical levels.

Subjects

The expert group consisted of 18 subjects who had completed at least 1 year of graduate study in physics. The novice group was composed of 18 undergraduates who had completed two semesters of coursework in physics at the college level. Novices with this level of physics training were selected to ensure that the novice group had had enough exposure to physics to have developed information-type rules, but had not approached the expert level in training. These subjects matched in level of expertise the experts and novices used in earlier research on novice importance rules (Dee-Lucas & Larkin, 1986, 1988).

Procedure

The subjects were told that the purpose of the experiment was to find out what types of information people find easy and difficult to understand in physics texts. They were told that they would read each passage three times and then complete a test on its contents. For the first reading, they were to read the passage to find out what it was about and its main points. The second and third times, they were to read more carefully in preparation for the exam. These instructions were given to reduce the potential variability in how subjects interpreted the study task.

The subjects were also told that the type of test that they would receive might vary for different passages. This was done so that readers would not have any strong expectations about what types of information they would be tested on, and so that the free recall test after the first target passage would not influence their reading strategy for the second target passage. To help maintain this expectation of varied test types, different types of tests were given after each of the practice passages. The goal was to force readers to rely on their own content schema for the domain of physics in deciding what was important to learn in the passages.

The passages were presented on a video display terminal one sentence at a time, with subjects controlling the sentence presentation rate by pressing a button. The study was conducted in two sessions. In each session, the subjects read one practice passage and one target passage. The order of presentation of the target passages was counterbalanced. After the first practice passage had been read three times, the subjects were asked to rate their overall familiarity with the passage content on a scale from 1 to 5, and then complete a multiple-choice test. After reading the target passage three times, the subjects again rated their overall familiarity with the passage topic and then were asked to recall the passage. When they had finished the free recall, they rated the individual target sentences on a 5-point scale according to how familiar they had been with the sentence content prior to reading the passage, and how easy it was to understand the sentence independent of its familiarity. The procedure for the second session was the same. However, the subjects were told that they would be given different types of tests in this session. After completing the readings of the practice passage, they completed a fill-in-the-blank test. After reading the target passage three times, they again completed a free recall test, and then rated the target sentences for their familiarity and comprehensibility.

RESULTS

Familiarity Ratings

An analysis of variance was performed on the familiarity ratings indicating the subjects' overall knowledge of the passage topics. The results indicate significant main effects for group [$F(1,35) = 48.37$, $MSe = 24.16$, $p < .001$] and passage [$F(1,35) = 12.76$, $MSe = 3.42$, $p < .001$]. The experts were more familiar with both passages ($M = 1.28$) than were the novices ($M = 2.42$). Both groups were more familiar with work and energy ($M = 1.65$) than fluid statics ($M = 2.08$).

Reading Times

The individual sentence reading times for definitions and facts were analyzed with a multiple linear regression performed on the log of the times expressed in milliseconds. A regression analysis was performed on individual sentence reading times so that each subject's familiarity and comprehensibility ratings for each sentence could be entered individually, and so differences in sentence length (i.e., number of propositions) could also be included. One data point was identified as an outlier through a normal probability plot of the residuals for the regression model, and this reading time was eliminated from the analysis. The resulting data set contained a total of 3,887 reading times.

The factors included in the analysis were passage (fluid statics, or work and energy), sentence type (definition or fact), level in the passage hierarchy (first, second, or third), comprehensibility and familiarity ratings, number of propositions, and reading trial (first, second, or third reading). Additionally, several interactions were coded in the analysis as the product of the corresponding main effect variables. The type × group and type × trial interactions were included in accordance with the study's hypotheses to see if interactions between these variables contributed to variance in the reading times. Also included

was the group × trial interaction to account for variance due to differences between groups in the degree to which they varied their overall reading speed across trials. The passage × group interaction was entered because preliminary analyses indicated that the main effect of passage was greater for one group than the other. Finally, the level × passage interaction was included because the magnitude of the effect of level varied with passage. Because the analysis was performed on the individual reading times for each subject on each target sentence, subjects were also included as a categorical variable. The regression coefficients for this model, excluding the subject variables, are presented in Table 1.

Sentence type effects. The positive coefficient for sentence type indicates that, averaged across reading trials, readers spent more time on definitions than on facts. However, the type × group interaction coefficient indicates that the size of this difference varied with level of expertise. Figure 1a shows the contributions of group and type to the estimated reading times for the first reading of the passages. These are the predicted reading times for the first reading trial for sentences consisting of eight propositions, which is the average length of the target sentences in this study.[3]

As can be seen in Figure 1a, the difference in the estimated times for definitions and facts is greater for novices than experts, with novices spending more time on definitions and less time on facts compared with experts. This is consistent with the type × group interaction found by Dee-Lucas and Larkin (1986) with importance judgments and summaries. It suggests that novice importance rules are reflected in attentional processes, in that novices differentiate between definitions and facts more than experts do in both judged importance and reading time.

The type × trial parameter indicates that the difference in the time spent on definitions and facts also varied with reading trial. The main effect of trial shows that the time spent on both definitions and facts decreased with each

Table 1
Parameter Estimates and Standard Errors for a Regression Model
of the Log of the Reading Times (in msec) on Definitions and Facts*

Variable	Coefficient	Standard Error	Standard Regression Coefficient
Intercept	3.510	—	—
Familiarity	.030	.004	.106
Comprehensibility	.027	.004	.095
Propositions	.037	.001	.418
Group	−.063	.021	−.104
Type	.072	.019	.120
Type × Group	.032	.014	.045
Trial	−.068	.007	−.184
Type × Trial	−.017†	.008	−.065
Group × Trial	.020	.008	.076
Level	−.026	.007	−.071
Level × Passage	.022	.009	.082
Passage	−.162	.019	−.268
Passage × Group	.060	.014	.086

*$R^2 = .52$; multiple $R = .72$.   †$p < .04$; all other coefficients significant at $p < .02$ level.
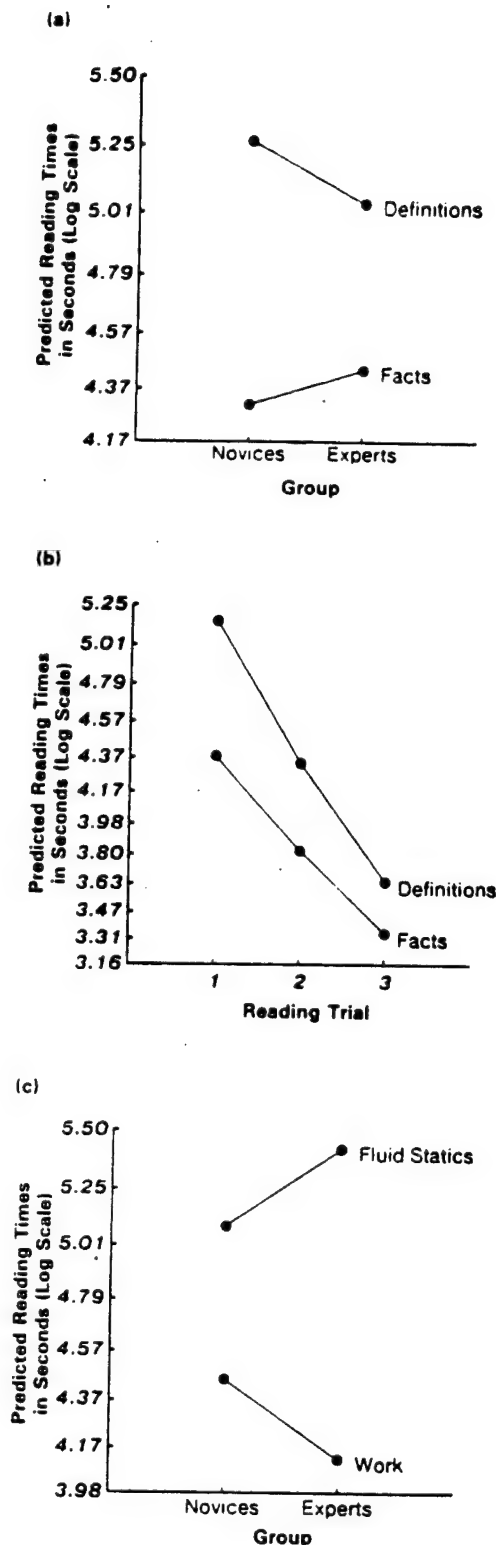
Figure 1. Estimated reading times for the initial reading of the passages, as determined by analyses of (a) group and sentence type, (b) reading trial and sentence type, and (c) passage and group.

rereading of the passage. However, the type × trial estimate shows that the decrease in reading time was greater for definitions than facts. This is shown in Figure 1b. This suggests that the greater processing time given to definitions than to facts decreased as the definitions were more thoroughly learned with each subsequent reading trial.

Group differences. Although experts read the target sentences more quickly than did novices, differences in reading speed between the two groups interacted with trial. The parameter for the group × trial interaction indicates that the decrease with each rereading (the main effect of trial) was more rapid for experts than for novices. Thus the difference between the reading times of the experts and the novices increased with each rereading because the experts decreased their reading times over trials more rapidly than did the novices. This may reflect greater confidence on the part of the experts that the subject matter had been sufficiently learned for the anticipated test.

Expert and novice reading times also varied with passage, as indicated by the positive parameter for the group × passage interaction. Readers in both groups spent more time on target sentences in the fluid statics passage (the less familiar passage) than in the work and energy passage. However, the group × passage estimate indicates that the difference between the amount of time spent on the target sentences in the two passages was greater for the experts than the novices. This interaction is shown in Figure 1c, which indicates that experts spent more time on fluid statics and less time on work and energy than did the novices. Thus the experts discriminated between the two passages to a greater degree than did the novices in terms of adjusting reading times.

Level effects. The parameter estimate for level indicates that readers spent more time on definitions and facts at the higher levels of the passage hierarchy, and less time on lower level sentences. However, the negative parameter for the level × passage interaction indicates that this effect was greater for the fluid statics passage than for the work and energy passage.

Other variables. Comprehensibility, familiarity, and number of propositions all had the expected effect on sentence reading times. The positive parameters for these variables indicated that readers spent more time on sentences that were less comprehensible, that were less familiar, and that contained more propositions. The fact that the comprehensibility and familiarity variables were significantly related to the variance in reading times suggests that subjects were able to adequately assess pre-experimental knowledge and syntactic complexity after reading the passages several times. The individual sentence comprehensibility and familiarity ratings correlated .18.

Free Recalls

The free recalls were scored at both the proposition level and the sentence level for the number of target sentences

and target sentence propositions recalled correctly and incorrectly. A correctly recalled sentence was defined as one in which all the core propositions were recalled accurately. Core propositions were considered to be the propositions expressing the main idea of the sentence. Partial credit was given for recall of sentences having more than one main idea. To receive partial credit, the subject had to recall all of the core propositions for at least one idea. A correctly recalled proposition was either one that was recalled verbatim or one that included substitutions that represented the gist of the proposition elements. Twelve of the 36 recalls were scored by a second scorer, and interscorer reliability was assessed. There was 95% agreement between the two scorers on the proposition scoring for the recalled sentences.

The free recalls were analyzed with an analysis of covariance (ANCOVA) performed on the mean number of target sentences recalled and on the proportion of total target sentence propositions recalled. The familiarity ratings were used as the covariate. This was done to control for recall differences due to prior familiarity with the sentence content. Analyses were performed on overall recall and on recall with errors excluded. All significance tests on the means were done using the Newman-Keuls test, with $p < .05$.

Proposition recall. An ANCOVA on the mean proportion of the total definition-related and fact-related propositions recalled at each level by each subject indicated significant main effects of type [$F(1,33) = 35.44$, $MSe = .01$, $p < .001$] and level [$F(2,67) = 68.82$, $MSe = .01$, $p < .001$], and a significant type × group interaction [$F(1,33) = 6.65$, $MSe = .01$, $p < .014$]. The type × group interaction is shown in Figure 2a. This interaction shows that the difference in the proportion of definition and fact propositions recalled is greater for novices than for experts. The adjusted cell means indicated that novices recalled .44 of the definition propositions and .30 of the fact propositions; experts recalled .48 of the definition propositions and .43 of the fact propositions. The difference in recall of definition and fact propositions is significant for the novices but not for the experts. Thus novices spent more time on definitions than on facts when reading the passages, and recalled significantly more definition- than fact-related information afterward. The difference between novice and expert recall was significant for facts but not for definitions.

The main effect of level is due to the fact that more target sentence propositions were recalled from the upper levels than from the lower levels of the passage hierarchies. All adjusted means were significantly different.

Propositions recalled accurately. Of an average total recall of 129 propositions from the target sentences, a mean of only 1.5 of these were recalled incorrectly. Because the number of recall errors was so low, the errors were not analyzed. Instead, the propositions recalled incorrectly were subtracted from the total recall data set, and the mean proportion of propositions recalled correctly by each subject were analyzed. The elimination of errors did not greatly alter the data set. The results of the anal-
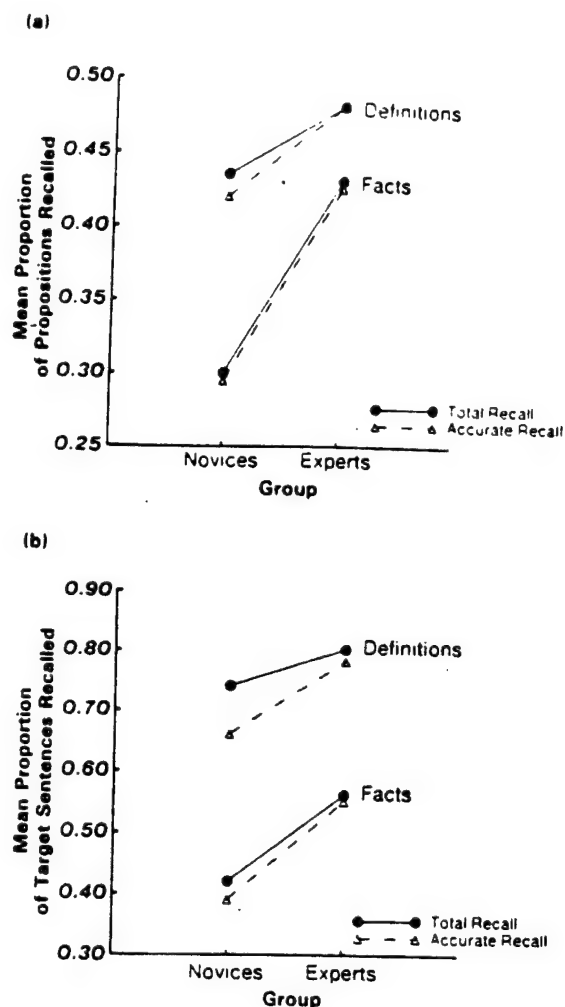
Figure 2. Mean proportion of (a) target-sentence propositions and (b) target sentences recalled by each group.

ysis indicated significant main effects due to type [$F(1,33) = 32.38$, $MSe = .01$, $p < .001$] and level [$F(2,67) = 65.51$, $MSe = .01$, $p < .001$], and significant interactions of type × group [$F(1,33) = 5.74$, $MSe = .01$, $p < .022$] and type × level [$F(2,67) = 3.06$, $MSe = .01$, $p < .053$]. The pattern of means for the type × group interaction was similar to that of the total data set (see Figure 2a). The type × level interaction is due to the fact that recall declined with level more sharply for facts than for definitions. The adjusted mean proportions of definition propositions recalled accurately were .52 (level 1), .47 (level 2), and .37 (level 3). For facts, these mean were .47, .40, and .22, respectively. All means were significantly different with the exception of definition recall at levels 1 and 2.

Sentence recall. An ANCOVA on the mean proportion of definitions and facts recalled from each level by each subject indicated significant main effects of type [$F(1,33) = 155.35$, $MSe = .03$, $p < .001$] and level [$F(2,67) = 95.96$, $MSe = .02$, $p < .001$], and a significant type × group interaction [$F(1,33) = 4.78$,

$MSe = .03, p < .036$]. These are the same effects found in overall proposition recall.

The type × group interaction shows that the difference in the proportion of definitions and facts recalled was greater for the novices than for the experts. The adjusted means for this interaction are plotted in Figure 2b. Novices recalled a mean of .74 definitions and .42 facts; experts recalled a mean of .80 definitions and .56 facts. Both groups recalled a significantly greater proportion of definitions than facts, but the novices recalled significantly fewer facts than the experts. There was not a significant difference between experts and novices in recall of definitions. This interaction again suggests that the subjects were most likely to recall the type of information that they spent the most time on during reading.

More high-level than low-level target sentences were recalled by both groups. All adjusted means for the three levels were significantly different.

Sentences recalled accurately. All sentences recalled incorrectly were eliminated from the total sentence recall data, and an ANCOVA was performed on the proportion of target sentences recalled correctly. Most of the recall errors occurred in core propositions. Eliminating errors at the sentence level therefore produced a greater alteration of the data set than did dropping errors from the proposition data, because the errors composed a larger proportion of the sentence recall data. However, the absolute number of errors was still very small. Subjects recalled an average of 22.7 of the 36 target sentences, and incorrectly recalled an average of 1.2 of these.

The analysis of the mean proportion of target sentences recalled correctly by each subject from each level indicated significant main effects of type [$F(1,33) = 104.43$, $MSe = .03, p < .001$] and level [$F(2,67) = 55.94$, $MSe = .03, p < .001$], and a significant type × level interaction [$F(2,67) = 5.74, MSe = .02, p < .005$]. There was not a significant type × group interaction in this data set [$F(1,33) = .71, MSe = .03, p < .41$].

The mean proportion of definitions and facts recalled accurately by novices and experts is plotted in Figure 2b for comparison with the type × group interaction found in the total sentence recall data. The greatest difference between the two data sets is in novice recall of definitions. The elimination of recall errors produced a more dramatic drop in the recall level for this cell than for any other. The fact that this did not occur with the proposition analyses indicates that novices did not recall more definition-related information incorrectly (relative to facts), but were more likely to recall incorrectly a core proposition from a definition. Thus the extra time spent on definitions resulted in more definitions being recalled, but not in more accurate recall. For a detailed analysis of the nature of the recall errors, see Dee-Lucas and Larkin (1987).

The type × level interaction is due to the fact that the greatest decline in definition recall occurs between levels 1 ($M = .86$) and 2 ($M = .67$; level 3, $M = .63$), whereas the largest decrease in fact recall occurs between levels 2 ($M = .50$) and 3 ($M = .28$; level 1, $M = .63$). There was no significant difference in the recall of definitions from levels 1 and 2; all other means were significantly different. As with the type × level interaction in proposition recall, the difference between the definition and fact recall is greatest at level 3.

DISCUSSION

Although novices, by definition, do not have fully developed knowledge structures for unfamiliar domains, recent research has shown that novices do form a rudimentary type of content schema for scientific texts. This schema is in the form of rules that specify what types of information are important in these texts. Only beginning-level students appear to possess these information-type rules; they are not found with naive subjects (i.e., those without domain-related training) or experts (Dee-Lucas & Larkin, 1988). Thus these rules are the result of limited domain-related training. The current study investigated the extent to which this type of domain-specific knowledge is used by novices in processing science texts. It also compared the general study strategies of novices and experts in adjusting reading time according to passage familiarity and structure.

Novice Rules and Text Processing

The results reported here indicate that novice importance rules do influence attention during reading. Novices consider definitions to be more important than facts when judging importance (Dee-Lucas & Larkin, 1986, 1988), and they spend more time on definitions than facts when studying physics texts. The use of this rule by novices was examined for both the initial reading and rereadings of the passages. The difference in reading times for definitions and facts was greatest on the first reading trial, indicating that these rules are "active" on the initial reading of the passage. Apparently, novices devote extra time to definitions as soon as they are identified during reading. This is consistent with research that indicates that readers automatically update their text representation with important text content as they read (Cirilo & Foss, 1980; Goetz, Reynolds, Schallert, & Radin, 1983; Just & Carpenter, 1980; Lorch, Lorch, & Matthews, 1985).

Novice importance rules also appear to influence learning: novices recall more definitions than facts. However, most of the novice recall errors occurred in recall of definitions. This suggests that novices may rate definitions higher in importance and spend more time on them because they find them difficult to learn, and not necessarily because they believe definitions are particularly important. However, the difference in recall accuracy between definitions and facts is very small. Moreover, experts also rate definitions as more important for novices to learn (Dee-Lucas & Larkin, 1986) and spend more time on definitions when they are reading physics texts, indicating that, on the average, definitions are more important for understanding physics. Additionally, novices rate definitions higher in importance even when the content of definitions and facts is identical, and therefore equally difficult

(Dee-Lucas & Larkin, 1988). Thus novice importance rules appear to partially reflect the nature of the content of physics, rather than to relate solely to learning difficulty.

Although there was a significant difference between recall of definitions and facts, recall differences may be somewhat attenuated because subjects read the passages three times. The subjects were able to concentrate on the facts and other types of information that they considered to be less critical on the second and third readings of the passages, possibly bringing the recall level for facts closer to that of definitions. The type × trial interaction in the reading time data suggests that subjects may have adopted this strategy. There was a greater decline in reading time over trials for definitions than for facts, indicating that the amount of attention given to facts increased relative to that given to definitions with subsequent passage readings.

The current study indicates that differences between the content schema of novices and that of experts are reflected in differences in the actual processing of domain-related texts. Thus novice rules are part of the control schema (Kintsch, 1982) used by novices to govern their text processing. The use of these importance rules during comprehension can be viewed as a "macrostrategy" within the framework of the Kintsch and van Dijk (1978; van Dijk & Kintsch, 1983) model of comprehension. Accordinging to this model, an integral part of the comprehension process is the formation of a "macrostructure" for the text (see also Guindon & Kintsch, 1984). The macrostructure is a representation of the gist or main ideas of the passage and is derived from the information in the text through the use of macrorules, which select from and reduce the text content. Macrostrategies are the sets of strategies used by readers in applying macrorules to construct a macrostructure. The use of novice information-type rules constitutes such a strategy in that these rules provide a basis on which to select the important text content.

Although this study indicates that novices use information-type importance rules as a macrostrategy in the naturalistic "study setting" provided, it is quite possible that they would use other rules under other circumstances. For example, under strong instructional sets, such as "learn all the formulas in this text," the influence of these rules would be attenuated or eliminated. However, these rules do appear to be part of the novice "default" strategy. That is, when novices are studying a physics text, they tend to expect certain types of information to be important for full understanding of that text. Moreover, these importance rules are very robust: they have been found in tasks requiring subjects to rate the importance of information, select the important content from texts, write summaries, and (in the current study) read and recall. They have also been replicated over many different novice groups (see Dee-Lucas & Larkin, 1986, 1988).

The use of these rules is a very sensible default strategy. As noted earlier, experts also judge definitions as more important than facts, and spend more time on definitions

than on facts. This suggests that definitions are typically important for understanding this content domain, and lack of attention to definitions may, in fact, result in comprehension difficulties. The problem with using this type of rule is that it is too general for accurately identifying what is important.

The experts and novices in this study were graduate students and undergraduates, respectively. These are the two subject groups that were used in previous research on novice importance rules. These groups were chosen in order to compare people who learn from physics texts (i.e., undergraduates) with people who use them in teaching (i.e., graduate students) to see if the groups match in terms of what they consider to be important to learn from physics texts (Dee-Lucas & Larkin, 1986). Because the current study attempted to replicate previous research findings using different dependent measures, the same subject groups were chosen. It is possible that the difference between these two groups in the time spent on definitions and facts was due to differences in their sensitivity to rhetorical indicators of importance, as opposed to differences in the perceived importance of different categories of information. That is, if the passages unintentionally signaled definitions as being more important than facts through the use of rhetorical devices, and if novices were more influenced than experts by these rhetorical signals during reading, then they would spend (relative to experts) more time on definitions and less time on facts. However, the fact that hierarchical structure had the same effect on the importance judgments of novices and experts suggests that the two groups did not differ in sensitivity to rhetorical indicators of importance. Additionally, previous research controlling the content and location of definitions and facts while varying only category membership shows that novices judge definitions as more important than facts when rhetorical indicators of importance are held constant (Dee-Lucas & Larkin, 1988). Furthermore, this earlier research compared the importance judgments of graduate students and undergraduates who were equivalent in physics background and found no difference between these two groups in the judged importance of definitions and facts, indicating that the definition-fact difference between experts and novices is due to differences in physics experience rather than differences in educational level.

Expert and Novice Reading Strategies

Although this study focused on the role of novice importance rules in text processing, it also permits a comparison of expert and novice strategies in rereading passages over several study trials. Both experts and novices appear to have adopted a "learning-to-criterion" strategy in varying their reading times according to sentence type and reading trial. There was very little difference between the groups in this respect. Subjects decreased the extra processing given to definitions over reading trials, so that the difference in the reading times for definitions and facts was very small by the third reading of the passage. Thus subjects appear to have concentrated on the most impor-

tant passage content on the first reading, and then gradually decreased the extra processing given to that content as it was more thoroughly learned. Although reading times decreased for facts as well, the decrease was not as rapid, so subjects were in effect increasing the amount of attention given to facts relative to that given to definitions with the passage rereadings. Although definitions are considered more important than facts, facts are also judged above average in importance (Dee-Lucas & Larkin, 1986, 1988), so it is reasonable to expect subjects to attend to both sentence types.

Although both experts and novices appear to have adopted similar strategies in learning the two sentence types, they differed in the extent to which they adjusted reading times according to reading trial and passage. Experts decreased their reading time more sharply over trials than did novices. This suggests that their greater expertise enabled them to learn the passage content more efficiently. This efficiency was also indicated by the fact that they recalled more target sentences overall than did novices, but read the target sentences more quickly. This is consistent with research that indicates that expertise facilitates learning from domain-related texts (Chiesi et al., 1979; Spilich et al., 1979). Spilich et al. found that baseball experts remembered more than novices after reading a passage describing a baseball game. They attributed the superior recall of the experts to their more elaborated and differentiated knowledge system, which allowed them to map more text content onto preexisting knowledge structures. The ease with which experts learned the content of physics texts in the current study is also most likely due to a large overlap between existing knowledge and text content. The experts could simply tag or elaborate existing knowledge structures in constructing a representation of the text, whereas novices had to generate new knowledge structures. Thus the experts could decrease reading time over trials more sharply than could novices because they did not need as thorough a review of the material in order to check that the appropriate knowledge had been tagged or to tag additional content. Novices, on the other hand, most likely used the rereadings to check, elaborate, and further generate new knowledge structures, tasks that are more time-consuming (see Johnson & Kieras, 1983).

Experts also distinguished between the two passages to a greater degree than did novices in adjusting their reading times. Although familiarity and comprehensibility differences at the sentence level were controlled through subject ratings, there was an additional effect of passage on the reading times. Both experts and novices spent more time on target sentences in one passage, but the experts spent more time on this passage and less time on the other passage than the novices. Because the two passages differ on a number of variables, the source of this effect is not clear. However, because the main effect of passage on reading time correlates with overall topic familiarity (with both groups spending more time on the less familiar pas-

sage), it suggests that general topic familiarity may have influenced reading times. If this is the case, then the passage × group interaction indicates that the experts were better able to adjust study time in accordance with passage topic, as well as reading trial. Thus the more elaborated knowledge structures of the experts may have allowed them to better determine how much study time was needed to master the content related to each topic.

This study controlled for passage structure in order to determine whether the influence of novice importance rules varied with hierarchical level. It was found that more time was spent on definitions than facts at all levels of the passage hierarchies. However, hierarchical level did have a main effect on readers' attention. Both experts and novices spent more time on high-level than low-level target sentences, and also recalled more high-level content. As discussed previously, novices are very adept at using various rhetorical indicators of importance, including passage structure (Kieras, 1985; Meyer, 1983). Thus, in the current study, novices were able to distinguish low-level elaborative facts and definitions from those that provided the main points of the passage. Therefore, novices apparently spent more time on definitions and less time on facts at all levels relative to the experts, with level modifying attention in a similar manner for both groups. This is consistent with the findings from the importance judgment research that indicate that novices judge high-level definitions and facts as more important than low-level content, but judge definitions overall higher in importance and facts overall lower in importance than do experts.

Although levels effects in reading time have been reported in other research (Cirilo & Foss, 1980; Schmalhofer & Glavanov, 1986), Schmalhofer and Glavanov found that this effect could be altered with instructional sets that implicitly indicated as important content lower in the passage structure. Thus levels effects in reading times appear to be a specific instance of the general principle that readers attend to the important text information, regardless of whether importance is defined on the basis of instructions, learning goals, text features, or schematic knowledge (see Anderson, 1982). This effect in reading times suggests that the influence of level on recall is due to selective attention: readers recognize high-level content as important, and therefore devote extra effort to learning it, resulting in longer reading times and better recall. However, this does not preclude alternative accounts of the levels effect in recall that do not rely on a selective attention hypothesis. The Kintsch and van Dijk (1978) model of comprehension attributes the effect to the repeated reprocessing received by high-level content as readers integrate low-level content with the high-level information. Additionally, Britton (Britton, Meyer, Hodge, & Glynn, 1980; Britton, Meyer, Simpson, Holdredge, & Curry, 1979) proposed several accounts of the levels effect which rely on retrieval mechanisms. It may well be the case that all of these phenomena contribute to levels effects in recall.

Conclusion

This research compared the effects of several variables on the attention and recall of novice and expert physicists, focusing specifically on sensitivity to information-type categories. The findings indicated that sentence category does influence the text processing of novices, with novices differentiating between category types to a greater degree than experts in adjusting their reading times. This is the same pattern found in importance judgments and summaries, and is also reflected in what novices remember from texts.

Given that novice importance rules influence attention and learning from text, the accuracy of these rules has important implications for the efficiency with which novices master text content. The fact that there is a partial mismatch between the content being learned by novices and that considered important for mastery by experts indicates that novices are attending to unimportant information (possibly learning this at the expense of other content) and missing some important content.

The findings from this study suggest that text writers for novice audiences will be most effective in enhancing learning if they use knowledge of subject-based content schema in conjunction with rhetorical devices in signaling the important text content. Writers need to know not only what is important in the text subject matter, but also what their readers think is important in order to be most effective in aiding learning of the critical text content.

REFERENCES

ANDERSON, R. C. (1982). Allocation of attention during reading. In A. Flammer & W. Kintsch (Eds.), *Discourse processing* (pp. 292-305). Amsterdam: North-Holland.

BRITTON, B. K., MEYER, B. J. F., HODGE, M. H., & GLYNN, S. (1980). Effects of the organization of text on memory: Tests of retrieval and response criterion hypotheses. *Journal of Experimental Psychology: Human Learning & Memory,* 6, 620-629.

BRITTON, B. K., MEYER, B. J. F., SIMPSON, R., HOLDREDGE, T. S., & CURRY, C. (1979). Effects of the organization of text on memory: Tests of two implications of a selective attention hypothesis. *Journal of Experimental Psychology: Human Learning & Memory,* 5, 496-506.

CHIESI. H., SPILICH, G., & Voss, J. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior,* 18, 257-274.

CIRILO, R., & Foss, D. (1980). Text structure and reading time for sentences. *Journal of Verbal Learning & Verbal Behavior,* 19, 96-109.

DEE-LUCAS, D., & LARKIN, J. H. (1986). Novice strategies for comprehending scientific texts. *Discourse Processes,* 9, 329-354.

DEE-LUCAS, D., & LARKIN, J. H. (1987). *Attentional strategies for studying scientific texts* (Tech. Rep. No. A.C.P. 17). Pittsburgh, PA: Carnegie-Mellon University, Department of Psychology.

DEE-LUCAS, D., & LARKIN, J. H. (1988). Novice rules for assessing importance in scientific texts. *Journal of Memory & Language,* 27, 288-308.

GOETZ, E. T., REYNOLDS, R. E., SCHALLERT, D. L., & RADIN, D. I. (1983). Reading in perspective: What real cops and pretend burglers look for in a story. *Journal of Educational Psychology,* 75, 500-510.

GRAESSER, A. C., HOFFMAN, N. L., & CLARK, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning & Verbal Behavior,* 19, 131-151.

GUINDON, R., & KINTSCH, W. (1984). Priming macropropositions: Evidence for the primacy of macropropositions in the memory for text. *Journal of Verbal Learning & Verbal Behavior,* 23, 508-518.

JOHNSON, W., & KIERAS, D. E. (1983). Representation-saving effects of prior knowledge in memory for simple technical prose. *Memory & Cognition,* 11, 456-466.

JUST, M. A., & CARPENTER, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review,* 87, 329-354

KIERAS, D. E. (1980). Initial mention as a signal to thematic content in technical passages. *Memory & Cognition,* 8, 345-353.

KIERAS, D. E. (1985). Thematic processes in the comprehension of technical prose. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 89-107). Hillsdale, NJ: Erlbaum.

KIERAS, D. E., & BOVAIR, S. (1981). *Strategies for abstracting main ideas from simple technical prose* (Tech. Rep. No. 9). Tucson: University of Arizona, Department of Psychology.

KINTSCH, W. (1974). *Representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

KINTSCH, W. (1982). Memory for text. In A. Flammer & W. Kintsch (Eds.), *Discourse processing* (pp. 186-204). Amsterdam: North Holland.

KINTSCH, W., & VAN DIJK, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review,* 85, 363-394.

LORCH, R. F., LORCH, E. P., & MATTHEWS, P. D. (1985). On-line processing of the topic structure of a text. *Journal of Memory & Language,* 24, 350-362. -

MEYER, B. J. F. (1975). *The organization of prose and the effect on recall.* Amsterdam: North-Holland.

MEYER, B. J. F. (1983). Text dimensions and cognitive processing. In H. Mandl, N. Stein, & T. Trabasso (Eds.), *Learning and understanding from text* (pp. 3-47). Hillsdale, NJ: Erlbaum.

SCHMALHOFER, F., & GLAVANOV, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representation. *Journal of Memory & Language,* 25, 279-294.

SPILICH, G. J., VESONDER, G. T., CHIESI, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior,* 18, 275-290.

VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

NOTES

1. The goal of Dee-Lucas and Larkin's (1986) research was to determine whether novices (i.e., physics students) agreed with experts (i.e., physics instructors) on what types of information are important for novices to learn in these texts. The experts were instructed to indicate the information in physics passages that they considered most important for their students to learn. Novices were told to indicate what they thought was most important to learn if they were going to be tested on the passage content.

2. The familiarity ratings were used to control for within-subject variations in prior knowledge of the definitions and facts. Johnson and Kieras (1983) showed that within-subject prior-knowledge differences measured by familiarity ratings are related to reading time. However, Johnson and Kieras collected familiarity ratings on the first reading of the passage and recorded reading times on the second passage reading. Because of the possibility of contaminating the reading time data for definitions and facts by having subjects read them first to rate their familiarity, the current study reversed this procedure and had subjects perform post hoc ratings of their preexperimental knowledge. This risked weakening the validity of the measure, in that subjects might judge the content as more familiar than it actually had been at the time of the initial reading. However, this would be indicated in the data analyses by a weak (or nonexistent) effect of familiarity on sentence reading times.

3. These reading times were calculated under the assumption that the sentences were very comprehensible and very familiar (i.e., familiarity and comprehensibility ratings of 1). This level of adjustment was selected for these two variables because the mean familiarity and comprehensibility ratings given to the target sentences indicate that subjects generally found the sentences to be easy to understand and familiar in content.

APPENDIX

The following is one of the passages used in this study. The target sentences are presented in bold, and the sentence type and hierarchical level are noted.

FLUID STATICS

In the preceding chapters, we studied the basic concepts needed to deal with systems consisting of many particles.

Hence, we can now discuss some of the important properties of common materials (such as solids, liquids and gases), which consist of many atoms or molecules.

We will spend most of our time discussing fluids (i.e., liquids and gases) because these have some remarkable properties of great importance for the understanding of physical and biological processes.

1. Pressure

Imagine a fluid in a container and in contact with some surface. Assume that the fluid and any other relevant objects (such as the container) are at rest so that the fluid is in equilibrium.

When this is the case, the fluid in contact with any surface exerts a contact force perpendicular to the surface and the surface exerts an equal and opposite force on the fluid. (fact, level 1)

These two forces must be perpendicular to the surface or the fluid layers will not be motionless.

The force exerted by a fluid on any object can be described by specifying the "pressure," which is defined as the magnitude of a fluid force divided by the area of the surface on which it acts. (definition, level 1)

This is a very useful concept for discussing fluid forces.

The amount of force produced by a fluid varies with surface area. (fact, level 2)

This force is exerted by molecules which are distributed over the entire surface exerting the pressure. (fact, level 3)

Therefore, if an area under pressure were made twice as big, for example, the number of molecules exerting a force would double and the total force exerted would become twice as large.

However, according to our definition, pressure is not proportional to area. If the area of the surface is made twice as big, the amount of pressure, which is equal to force divided by area, remains the same.

Although all fluids exert pressure, the pressure of the air in the atmosphere near the earth's surface is of special interest because we usually work in an environment of air at atmospheric pressure.

We can define atmospheric pressure at any point as the ratio of the weight of a column of air extending from that point to the top of the atmosphere, divided by the cross-sectional area of that column. (definition, level 2)

According to our definition, atmospheric pressure decreases with increases in altitude as the weight of the air above that point decreases. *Atmospheric pressure also varies from day to day due to changes in air temperature.* (fact, level 3)

Pressure is often expressed in terms of "gauge pressure," which is defined as the difference between the pressure at a point in a fluid and the atmospheric pressure. (definition, level 2)

The pressure at a point in a fluid is called the "absolute pressure," and is defined as simply the actual pressure at that point. (definition, level 3)

Since we live and work in an environment at atmospheric pressure, gauge pressure uses atmospheric pressure as a standard against which the actual pressure is compared.

For example, a mercury-filled manometer is commonly used to measure the difference between arterial blood pressure and the atmospheric pressure.

The measured blood pressure is then the gauge pressure of the blood. According to our definition of gauge pressure, the gauge pressure is negative when the absolute pressure is less than the standard atmospheric pressure.

2. Density

Any small portion of a material has some volume and some mass (which is the sum of the masses of all the atoms contained in this portion). *If this portion is small enough, the number of atoms in it, and thus also its mass, is proportional to its volume.* (fact, level 1)

For example, if the volume of a portion of water was increased so that it was 3 times as large, the number of atoms in this portion, and thus also the mass of this portion, would be 3 times as large.

We can specify a ratio called "density," which is defined as the mass of a portion of material divided by its volume, or mass per unit volume. (definition, level 1)

The density of a substance does not change with changes in volume. (fact, level 2)

Properties with this characteristic are called "intrinsic" quantities, which are defined as quantities which are independent of volume. (definition, level 3)

The density of a substance often varies with the position of the portion of the material taken, since there are often variations in the internal characteristics of substances. (fact, level 2)

However, density does not always depend on position of material sampled.

Some materials are "homogeneous," which is defined to mean that all intrinsic properties of the material are the same for every portion of the material. (definition, level 3)

Thus, in materials which are homogeneous, the density is the same at each point.

It is sometimes convenient to re-express density in terms of specific gravity.

Specific gravity is defined as the ratio of the density of a substance to the density of water. (definition, level 2)

For example, pure water has a specific gravity of exactly 1, while mercury has a specific gravity of 13.6.

The specific gravity of a substance varies with its temperature. (fact, level 3)

Therefore, a specific gravity value for a substance is accompanied by a statement of the temperature at which it was evaluated.

3. Buoyant Force

Suppose an object is surrounded by a liquid at rest, and this object, having a density less than the density of the liquid, floats partly submerged in this liquid.

The pressure exerted by the fluid on each part of the surface of this submerged body does not depend on the material that the body is made of. (fact, level 1)

If this hypothetical object is replaced by fluid similar to that of its surroundings, for example, this body of fluid will experience the same pressures that acted on the immersed object, and will be at rest.

An object floating or submerged in a liquid experiences a "buoyant force" due to the surrounding liquid, which is defined as the product of the density of the liquid, the gravitational acceleration g, and the volume V, of the liquid displaced by the object. (definition, level 1)

Thus the buoyant force is simply the total force exerted on the object by the surrounding fluid.

<u>Appendix B</u>

(Total pages in length: 21)

Dee-Lucas, D. & Larkin, J.H. (1988b)

Novice rules for assessing importance in scientific texts.

Journal of Memory and Language, **27**, 288-308

Novice Rules for Assessing Importance in Scientific Texts

Diana Dee-Lucas and Jill H. Larkin

Carnegie Mellon University

Previous research shows that content area novices judge certain categories of information (e.g., definitions, facts, equations) as more important than others. The current research tested the hypothesis that novice importance judgements are based on category membership, rather than content differences between categories. Subjects of varying expertise judged the importance of sentences in physics texts when they were presented in one of two forms: definitions or facts (Experiment 1), and equations or their verbal equivalents (Experiment 2). The two sentence versions were always identical in substantive content. Experts and naive subjects (subjects without physics training) judged these variants to be similar in importance. However, beginning physics students judged definition and equation versions as more important. Thus beginning-level students develop rules specifying what categories of information are important, so that sentence category is a salient text feature. Sentence category is irrelevant for experts, who judge importance according to content, and naive subjects, who have not formed expectations regarding the importance of information categories. These results suggest how a content schema might evolve in novice learners. © 1988 Academic Press, Inc.

Scientific textbooks are typically densely packed with complex information, including equations, symbols, and specialized terms. Consequently, it can be difficult for students who are unfamiliar with scientific subject matter to distinguish the important content from the elaborative information when reading this type of text. There are various information sources that novice readers (i.e., readers who are unfamiliar with the text content domain) could use in assessing importance. Most research has focused on text-based indicators of importance such as text structure and signaling devices (e.g., underlining, adjunct questions, staging, typographical cuing). These textual manipulations are "content-

free," in that their effects should not depend on the nature of the text content or the expertise of the reader. In contrast, a "content-specific" source of information for assessing importance in texts is the reader's "content schema" (Kieras, 1985). A content schema consists of domain-specific knowledge about how information in a content area is typically organized, including what is important.

Recent work by Dee-Lucas and Larkin (1986) suggests that novices develop a rudimentary content schema for scientific domains consisting of rules specifying what types of easily-recognizable information (i.e., definitions, facts, equations) are important. The purpose of the present research was to investigate the nature of these rules. The studies reported here tested the hypothesis that novices judge importance on the basis of category-membership rules—that is, that they consider certain information to be important simply because of its category membership (i.e., whether it is a definition, equation, fact), regardless of its content. This was done by varying the category membership of se-

lected information in physics texts, and examining how the category changes influenced experts' and novices' judgements of the importance of the content. Two experiments contrasted the judged importance of identical information when it was presented in different forms; specifically, when it was stated as a definition or a fact, and written as an equation or in sentence form.

Novice importance rules have been investigated in research comparing the importance judgements of expert and novice physicists for different types of information in physics texts (Dee-Lucas & Larkin, 1986). Although this research found that experts and novices generally agreed on the relative importance of various types of information, novices' importance judgements were more strongly related to type categories than were experts' judgements. For example, both groups judged definitions to be more important than facts, but novices were even more likely than experts to judge definitions as important and facts as unimportant. This suggests that the novices had formulated a general rule that definitions are more important than facts and judged importance on this basis. Unlike novices, the experts' importance judgements were not tied as closely to category membership. They presumably were judging importance on the basis of the nature of the specific sentence content, rather than the type category.

The findings from this area of research suggest that people just beginning to learn about physics develop a set of rules defining what types of information are important in that domain. They use these rules both in deciding what is important (Dee-Lucas & Larkin, 1986) and in guiding attention during reading (Dee-Lucas & Larkin, in press). However, this research on novice importance rules was based on naturalistic materials, and thus did not control for content differences between information categories. Therefore, novices in these studies may have been basing their importance judgements on some feature of the *content* that differed between the categories, rather than entirely on the categories themselves. For example, novices may have considered definitions to be more important than facts because the definitions contained more new terms. If this were the case, then the rule used by the novices would be that statements with new terms are more important than statements involving known terms, rather than a rule that definitions are more important than facts. Thus it is not known whether novices are judging importance according to a superficial analysis of category membership, or a "deeper" analysis of the nature of the information typically contained in various categories. A more precise understanding of the basis for novice importance rules would suggest how domain-specific content schema evolve in novice learners.

The current research tested the hypothesis that novice importance rules are based on category membership. This was done by varying the category membership of specific information in physics texts, and examining how the category change influenced experts' and novices' judgements of the importance of the content. If novice importance rules are based on category membership, then novices should vary their assessment of the importance of a given statement with changes in its category. Experts, on the other hand, should be relatively uninfluenced by category changes because they would presumably judge importance on the basis of the nature of the content. The category membership of sentence information was manipulated through minor wording variations which did not alter the primary sentence content, but signaled the content as belonging to a particular category. In this way, category membership could be varied by changing the form in which the information was presented while content was held constant.

Two studies were conducted in which expert and novice physicists judged the importance of information in physics texts.

Experiment 1 compared the judged importance of attributive information when it was presented in the form of a definition or a fact. The purpose was to determine whether differences in the perceived importance of definitions and facts found in earlier research were due to differences in category membership as opposed to content differences. This was determined by contrasting the judged importance of information when it was signaled as being a definition, through the use of the words "is defined as," and when this signaling was absent or replaced with a neutral phrase so that the content appeared to be a fact. In this way the judged importance of definitions and facts was compared while holding sentence content constant across categories.

Experiment 2 extended the findings from the first study to a class of information which is particularly important in physics problem-solving—quantitative relations. It contrasted the judged importance of quantitative relations when they were presented as equations (e.g., $a = b/c$) or written out in verbal form (e.g., a is equal to b divided by c). The equational form signals the content as being quantitative in nature, while the verbal form makes the category membership of the content less apparent. Previous research indicates that novices (as well as experts) consider equations to be a particularly important type of content in physics texts (Dee-Lucas & Larkin, 1986). If novices are basing this judgement on a category-membership rule, (i.e., that equations are important regardless of content) then they would judge the same quantitative information as more important when it was presented as an equation rather than in verbal form.

In each study, two physics passages were used which contained target sentences that could be expressed in different forms. There were two versions of each passage, each version being identical except for the form in which the target sentences were presented. The importance

judgements of the novices and experts for the fact and definition versions (Experiment 1) and the equation and verbal versions (Experiment 2) of the target sentences were compared to see to what extent sentence form influenced the perceived importance of the information.

EXPERIMENT I

The goal of this experiment was to determine whether novices consider definitions to be more important than facts independent of the content of the two sentence types. It also clarified why experts consider definitions to be more important than facts in natural texts. This effect (though not as large as for novices) could be due to category-type rules, or because the content of definitions is typically more important than that of facts. If experts are judging importance on the basis of content, then they should be relatively uninfluenced by changes in the category membership of that content.

Method

Stimulus materials. One passage was about work and energy and one dealt with fluid statics. Each was about 50 sentences long. One contained 9 target sentences and one had 11 target sentences.

The definition and fact versions of the target sentences differed in that definitions always included "is defined as," and thus were signaled as being definitions. In the fact versions, "is defined as" was dropped or replaced with "is represented as," "is calculated as," or "is indicated by." Thus the facts were "nondefinitions" in that in place of definition signaling they contained phrases indicating that the sentence was presenting attributive information about the sentence topic (as opposed to criterial attributes defining the sentence topic).

Examples of the definition and fact versions of some of the target sentences are shown in Table 1a. There were two versions of each passage. In one version, the odd-numbered target sentences were defi-

TABLE 1
SAMPLE TARGET SENTENCES

(a) Examples of definition and fact versions (Experiment 1)

1. Absolute pressure is defined as simply the actual pressure at a point.

   Absolute pressure is simply the actual pressure at a point.

2. Specific gravity is defined as the ratio of the density of a substance to the density of water.

   Specific gravity is indicated by the ratio of the density of a substance to the density of water.

3. In terms of this notation. the work ΔW done by a force F in moving an object through a displacement Δr is defined as

$$\Delta W = \mathbf{F} \cdot \Delta \mathbf{r} = F(\Delta r \cos \theta).$$

   In terms of this notation. the work ΔW done by a force F in moving an object through a displacement Δr is represented as

$$\Delta W = \mathbf{F} \cdot \Delta \mathbf{r} = F(\Delta r \cos \theta).$$

4. The unit typically used for measuring work. the joule. is defined as the work done by a unit force (one newton) acting on a unit distance (one meter).

   The unit typically used for measuring work. the joule. indicates the amount of work done by a unit force (one newton) acting on a unit distance (one meter).

5. Pressure is defined as the magnitude of a fluid force divided by the area of the surface on which it acts.

   Pressure can be calculated by dividing the magnitude of a fluid force by the area of the surface on which it acts.

(b) Examples of equation and verbal versions (Experiment 2)

1. Kinetic energy K is equal to the product of one-half the mass m of the particle times the square v^2 of its speed.

   Kinetic energy is

$$K = 1/2 \, mv^2.$$

where m is the mass of the particle and v is its speed.

2. The velocity v of an object is equal to the rate at which its position changes with time. or the displacement Δr divided by the corresponding time interval Δt.

   The velocity of an object is equal to the rate at which its position changes with time or

$$v = \Delta r/\Delta t.$$

where Δr is the displacement and Δt is the corresponding time interval.

3. The gauge pressure P_g at a point in a fluid is equal to the difference between the pressure P_p at that point and the atmospheric pressure P_{atmos}.

   The gauge pressure at a point in a fluid is

$$P_g = P_p - P_{atmos}.$$

where P_p is the pressure at that point and P_{atmos} is the atmospheric pressure.

4. Density ρ is equal to the mass m of a portion of material divided by the volume V of that portion.
   Density is

$$\rho = m/V.$$

where m is the mass of a portion of material and V is the volume of that portion.

5. Specific gravity S is equal to the density ρ_s of a substance divided by the density ρ_w of water.
   Specific gravity is

$$S = \rho_s/\rho_w.$$

where ρ_s is the density of a substance and ρ_w is the density of water.

---

nitions and the even-numbered were facts: in the second version this was reversed.

Each of the target sentences was classified according to its level in the hierarchical structure of the passage. The procedure used for the structural analysis is reported in Dee-Lucas and Larkin (1986, in press). This analysis produced a hierarchy with the main topics or concepts occurring at the highest levels and modifying information occurring at the lower levels. Modifying information consisted of examples. attributes and properties. derivations (i.e.. information implied by or derived from higher level information). explanations. subtopics. and preconditions (i.e.. necessary conditions for a rule. principle. or fact to hold true). The hierarchical analysis was performed at

the sentence level. There were 7 sentences at level 1 (the most superordinate level), 6 sentences at level 2, and 7 sentences at level 3. Hierarchical level was included as a variable in the data analyses to see whether perceived importance was influenced by level, and if this variable interacted with sentence form (i.e., definition or fact).

Subjects. The novices were 24 undergraduates with 2 or 3 semesters of college physics. Novices with this level of physics training were selected to ensure that the novice group had had enough exposure to physics to have developed information-type rules, but had not approached the expert level in training. The 24 experts had completed at least one year of graduate study in physics.

Two control groups were also run in the experiment to see if expert–novice differences in the perceived importance of the target sentences were due to differences in educational level (i.e., undergraduate vs. graduate level training) as opposed to differences in physics knowledge. The two control groups were selected so as to differ in their educational level in the same manner as the two experimental groups. However, none of the control group subjects had taken any college-level physics, so they were similar to each other in terms of their physics knowledge. The undergraduate control group consisted of 24 undergraduates; the graduate student control group consisted of people who had completed at least 1 year of graduate training in the humanities or social sciences. Although this group will be referred to as the graduate student control, it included some postdoctoral researchers and faculty. This was also true of the corresponding expert experimental group.

Although the control groups were specifically selected to control for educational differences, they would also indicate differences in perceived importance due to age, maturity, and verbal ability. In the case of verbal ability, it is reasonable to assume that graduate students in the social sciences/humanities would be as high in verbal ability as graduate students in physics. Similarly, it is likely that undergraduates attending the same university are roughly equivalent in verbal ability.

Procedure. The subjects were given one version of each passage. They were told to read each passage carefully, then rate the importance of each sentence on a scale from 1 (most important) to 5 (least important), and then indicate the 10 most important sentences in each passage. The instructions for the rating task indicated that each rating should be used at least once. All of the sentences were rated, but only the ratings for the target sentences were analyzed. The order in which the passages were read and the versions of the passages received were counterbalanced.

The novice and the undergraduate control groups were told that in completing the tasks, they were to indicate which sentences they thought would be most important to learn if they were going to be tested on the passage content. The expert and the graduate student control groups were told to pretend that they were teaching a course and indicate which sentences they thought were most important for their students to learn. These instructions match those used by Dee-Lucas and Larkin (1986) in their initial research on expert–novice differences in perceived importance. The instructions were designed to compare what novices think they should learn with what experts (their instructors) think novices should learn.

Results

The data from the two dependent measures were analyzed in several ways. The data from the ratings task were set up as a frequency table and analyzed using a multiway frequency analysis. This analysis fits a loglinear model to categorical data. The loglinear model expresses the logarithm of the expected cell frequencies as an additive function of main effects and interactions, in a manner similar to the typical analysis of

variance model (Fienberg, 1980). For this analysis, the number of responses in each rating category (1 through 5) for each sentence type (definition and fact) occurring at each level (1 through 3) was tabulated for each of the subject groups. The multiway frequency analysis was performed on the total number of responses occurring in each of these cells.[1]

The data from the sentence selection task (i.e., select the 10 most important sentences) consisted of an indication for each subject of whether or not each target sentence was selected as one of the most important.[2] Because the dependent variable (selected/not selected) is dichotomous, the data were analyzed using a logistic regression (Fienberg, 1980). The variables included in the analysis were sentence type (definition or fact), level (1 through 3), and subject group (novice or expert).

The data from the two control groups were submitted to identical analyses. The results of these analyses were compared to the results of the corresponding analyses of the experimental group data to determine if expert–novice differences were also reflected in differences between graduate students and undergraduates who had had no advanced physics training.

The results from all of these analyses have been presented together in Fig. 1 in order to show the consistency of findings across dependent measures. Thus this figure is referred to repeatedly in the discussion of the results.

In addition to the analyses discussed above, a $2 \times 2 \times 3$ analysis of variance

was performed on each data set, with group as a between subjects variable and sentence type and hierarchical level as within subjects variables. For the ratings data, this analysis was performed on the mean rating given to each sentence type at each level by each subject. For the sentence selection task, the analysis was performed on the mean proportion of sentences of each type selected as important from the three levels by each subject. The results of these analyses will be discussed only as they pertain to the main finding of interest, the type $\times$ group interaction. The means for this interaction for the two dependent measures are presented in Table 2.

Ratings data: Experimental groups. The multiway frequency analysis of the ratings data indicated that the best-fitting model was a hierarchical model including the type $\times$ group interaction and the main effect of level ($\chi^2 = 23.80$, $df = 30$, $p < .78$). The ANOVA also indicated a significant type $\times$ group interaction ($F(1,46) = 6.95$, $p < .01$). The mean ratings predicted by the loglinear model for the type $\times$ group interaction are shown in Fig. 1a. The predicted means for the novices are 1.67 when the sentence was in the form of a definition and 1.89 when it was in the form of a fact. For experts, the predicted ratings are 1.82 for definitions and 1.79 for facts. The actual means for this interaction are shown in Table 2. This interaction indicates that

[1] There were 480 observations for both the novice group and the undergraduate control group. In the expert group, one subject overlooked two sentences in performing the ratings, resulting in a total of 478 observations for this group. In the graduate control group, one sentence was overlooked, resulting in 479 observations for this group.

[2] It was assumed that the sentences that subjects missed in performing the sentence ratings were also overlooked in the sentence selection task, and these three data points were excluded from the analysis.
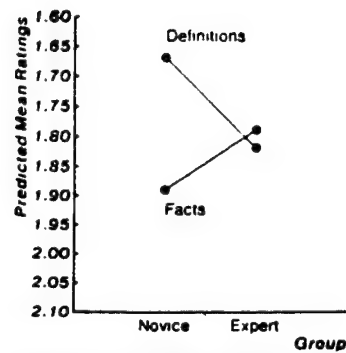
TABLE 2

MEAN IMPORTANCE RATINGS AND MEAN PROPORTION OF TARGET SENTENCES SELECTED AS IMPORTANT FOR THE TYPE × GROUP INTERACTIONS (EXPERIMENT 1)

	Experimental groups		Control groups	
	Novice	Expert	Undergrad	Grad
Ratings				
Definition	1.66	1.81	1.52	1.95
Fact	1.90	1.76	1.65	2.04
Proportion selected				
Definition	.65	.51	.61	.48
Fact	.45	.49	.54	.42

(a) Predicted mean ratings for the type x group interaction -- experimental groups.

(b) Predicted mean ratings for the type x group interaction -- control groups.
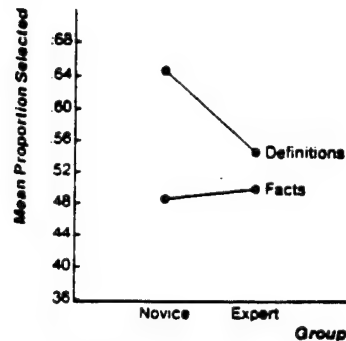
(c) Predicted proportion of target sentences selected as important for the type x group interaction -- experimental groups.
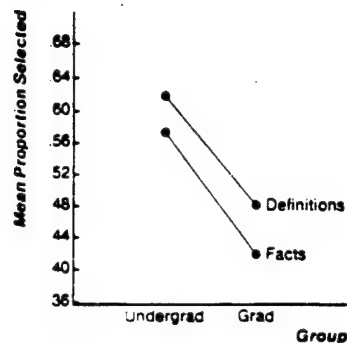
(d) Predicted proportion of target sentences selected as important for the type x group interaction -- control groups.

(e) Predicted proportion of target sentences selected as important for the type x level interaction -- control groups.

FIG. 1. Comparisons of the judged importance of definitions and facts (Experiment 1).

novices were influenced in their importance ratings by the form in which the information was presented. They considered the same content to be more important if it was stated as a definition as opposed to a fact. The experts, on the other hand, did not appear to base their ratings on sentence form; there is very little difference in their predicted mean ratings for definition and fact versions of the target sentences.

The parameter estimates for the main effects and interaction for the complete

TABLE 3
Parameter Estimates. Standard Errors. and Ratio of Estimates to Standard Errors for a Loglinear Model of the Ratings Data from the Experimental Groups (Experiment 1)

Effect	Coeff.	St. Error	Ratio: Coeff. St. Error
(a) Group: Novice estimates (expert estimates are opposite)			
Rate 1	.045	.066	.69
Rate 2	.074	.075	.98
Rate 3	.094	.090	1.04
Rate 4	.019	.118	.16
Rate 5	-.232	.182	-1.27
(b) Type: Definition estimates (fact estimates are opposite)			
Rate 1	.073	.067	1.10
Rate 2	.082	.075	1.10
Rate 3	-.224*	.090	-2.48
Rate 4	.064	.115	.56
Rate 5	.005	.192	.03
(c) Type × Group: Definition estimates (fact estimates are opposite)[a]			
Rate 1			
Novice (+)	.133*	.066	2.01
Expert (−)			
Rate 2			
Novice (−)	.098	.075	1.30
Expert (+)			
Rate 3			
Novice (+)	.003	.081	.04
Expert (−)			
Rate 4			
Novice (+)	.014	.117	.12
Expert (−)			
Rate 5			
Novice (−)	.052	.180	.29
Expert (+)			
(d) Level			
Level 1			
Rate 1	.917*	.121	7.57
Rate 2	.214	.135	1.59
Rate 3	-.454*	.185	-2.46
Rate 4	-.504*	.250	-2.01
Rate 5	-.173	.351	-.49
Level 2			
Rate 1	-.042	.110	-.38
Rate 2	.052	.120	.44
Rate 3	.140	.148	.95
Rate 4	.165	.193	.86
Rate 5	-.315	.320	-.98
Level 3			
Rate 1	-.875*	.098	-8.97
Rate 2	-.265*	.104	-2.56
Rate 3	.314*	.128	2.46
Rate 4	.339*	.169	2.00
Rate 5	.488	.251	1.94

Note. Asterisks indicate coefficients that differ from zero by more than 2 standard errors.

[a] Parameter estimates indicate the size of the effect: positive and negative symbols indicate the direction of the effect for each group.

model are shown in Table 3. Because of the usual constraints placed on the model. all parameter estimates for each main effect and interaction are constrained to sum to zero. Therefore. for all effects the magnitude of the parameter estimates for each variable are the same but in the opposite direction. The ratios of the estimates to the standard errors indicate the degree to which the parameter estimates differ from zero.

The parameter estimates for the main effect of group show little difference between the experts and novices in their use of the five ratings categories. The largest differences occurred in the use of rating 5 (the lowest rating) and rating 3 (the middle rating). The novices tended to use the rating 3 category more often than the experts (as indicated by the positive parameter estimate). while the experts tended to use the lowest rating more often than the novices. This suggests that the experts rated the target sentences lower in importance than the novices.

The parameter estimates for the main effect of type indicate that the largest difference occurred in the rating 3 category (the middle rating). The negative parameter estimate indicates that this rating was used more often with facts than definitions. There were also smaller differences in the use of the first two rating categories. with the definitions rated 1 or 2 more often relative to facts. This indicates that definitions were rated higher overall in importance than facts.

The type × group interaction estimates indicate that the greatest differences between experts and novices in rating definitions and facts occurred in the first two rating categories. The novices were more likely than experts to give a target sentence a rating of 1 if it was in the form of a definition. and somewhat more likely to rate it a 2 if it was in the form of a fact. The opposite was true for the experts. relative to the novices.

Level had a very strong influence on the target sentence ratings. as shown by the large parameter estimates for this effect. Level 1 target sentences tended to be rated as most important. indicated by the large

positive estimate for the rating 1 category.
The ratings for level 2 target sentences
were spread over the categories without
any strong clustering in any one rating:
none of the parameter estimates for level 2
sentences differed from zero by more than
two standard errors. Level 3 target sen-
tences tended to be rated as 3 or 4 in im-
portance, indicated by the positive param-
eters for these ratings categories. Level did
not interact with type or group in in-
fluencing the ratings (i.e.. including these
interactions reduced the fit of the model).

Ratings data: Control groups. The mul-
tiway frequency analysis of the control
group data indicated that the best-fitting
model included only the main effects of
group and level ($\chi^2 = 28.64$, $df = 40$, $p <$
.91). The inclusion of the type × group in-
teraction or the main effect of type reduced
the fit of the model to the data set. The
type × group interaction was also not sig-
nificant in the ANOVA ($F(1.46) = .24$, $p <$
.63). The mean ratings predicted by the
loglinear model with the type × group in-
teraction included are presented in Fig. 1b
for comparison with the corresponding ex-
perimental group means. The actual means
are presented in Table 2. As Fig. 1b shows,
there was no difference between the under-
graduate and graduate controls in the influ-
ence of sentence type on the mean ratings
of the target sentences. Additionally, the
lack of a main effect of sentence type indi-
cates that the form in which the target sen-
tences were presented did not influence the
control group ratings (i.e., they did not
consider definitions to be more important
than facts).

The parameter estimates for the loglinear
model including the main effects of group
and level are shown in Table 4. The esti-
mates for the main effect of group indicate
that the undergraduates and graduates dif-
fered primarily in their use of ratings 1 and
4. The positive estimate for the undergrad-
uates for rating category 1 indicates that
they rated target sentences 1 more often

than the graduate controls. The opposite
was true for the rating 4 category. with the
graduates using this rating more often rela-
tive to the undergraduates. This indicates
that undergraduates rated the sentences as
more important overall than the graduate
controls. This effect is also apparent in
Fig. 1b.

The pattern of parameter estimates for
the main effect of level is similar to that ob-
tained with the experimental groups. Level
1 target sentences were most likely to re-
ceive a rating of 1. indicated by the large
positive parameter estimate for that rating.

TABLE 4

PARAMETER ESTIMATES. STANDARD ERRORS. AND
RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A
LOGLINEAR MODEL OF THE RATINGS DATA FROM
THE CONTROL GROUPS (EXPERIMENT 1)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
(a) Group: Undergrad estimates (grad estimates are opposite)			
Rate 1	.323*	.065	4.98
Rate 2	.083	.074	1.12
Rate 3	−.070	.089	−.79
Rate 4	−.323*	.117	−2.76
Rate 5	−.010	.172	−.06
(b) Level			
Level 1			
Rate 1	.695*	.105	6.62
Rate 2	.141	.119	1.18
Rate 3	−.302	.156	−1.94
Rate 4	−.492*	.220	−2.24
Rate 5	−.042	.282	−.15
Level 2			
Rate 1	.128	.102	1.25
Rate 2	.013	.118	.11
Rate 3	.133	.137	.97
Rate 4	−.047	.188	−.25
Rate 5	−.226	.281	−.80
Level 3			
Rate 1	−.822*	.092	−8.92
Rate 2	−.154	.099	−1.56
Rate 3	.169	.119	1.42
Rate 4	.539*	.155	3.48
Rate 5	.268	.226	1.19

Note. Asterisks indicate coefficients that differ from
zero by more than 2 standard errors.

The ratings given to level 2 target sentences were spread over the categories. with the strongest clustering in the rating 1 category (though this parameter estimate was much smaller than the rating 1 estimate for level 1 sentences). Level 3 target sentences were most likely to receive a rating of 4. with 3 as the next most frequent rating category for this level.

Ratings data: Summary. The results of the multiway frequency analyses indicate that novices base their judgements of the importance of text information on the category membership of the content. They specifically rate the same attributive information as more important when it is presented as a definition as opposed to a fact. as is shown in Fig. 1a. Expert physicists are not influenced by category in rating importance. presumably basing their importance judgements on the nature of the sentence content. Similarly. the physics-naive control subjects are not influenced by signaled category membership. so that they appear to behave like experts in rating importance. This can be seen in Fig. 1b. This is most likely because these subjects have no strong expectations about the relative importance of definitions and facts in physics texts. and thus are not influenced by this text feature. The lack of a type × group interaction in the control group data indicates that expert–novice differences in the perceived importance of definitions and facts are not due to differences in educational level.

Sentence selection data: Experimental groups. The sentence selection data were analyzed using a logistic regression. The regression analysis indicated that a good fit to the experimental group data was provided by a hierarchical model including the group × type interaction and the main effect of level ($\chi^2 = 8.69$. $df = 7$, $p < .28$). The regression model is the same model found to provide the best fit for ratings data from the experimental g   . The ANOVA also indicated a signifi... type × group interaction ($F(1.46) = 10.01$, $p <$

.01). The predicted mean proportion of target sentences selected for the type × group interaction are shown in Fig. 1c plotted on a logit scale. The actual means are presented in Table 2. This interaction is very similar to the type × group interaction obtained with the ratings data. It shows that the novices were more likely to select a target sentence as important when it was presented in the form of a definition than a fact. while the experts were relatively unaffected by sentence form in their selection of the important sentences.

The parameter estimates for the complete logistic regression model are shown in Table 5. Unlike the estimates for the multiway frequency analysis. the logistic regression estimates show the size of the difference between the means for the two variables in an effect. Therefore only one parameter estimate is presented for each main effect (two for the interaction) and the ratios indicate the size of the difference between the two variables in each effect and interaction.

The negative estimate for the main effect of type indicates that subjects tended to select more target sentences when they were presented in the form of a definition. This is consistent with the main effect of type found with the ratings data from the experimental groups. The estimate for the main effect of group indicates little difference be-

TABLE 5
PARAMETER ESTIMATES. STANDARD ERRORS. AND RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A LOGISTIC REGRESSION MODEL OF THE SENTENCE SELECTION DATA FROM THE EXPERIMENTAL GROUPS (EXPERIMENT 1)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
Type	−.414	.099	−4.18
Group	−.186	.099	−1.88
Type × Group Definition (−) Fact (+)	.236	.099	2.38
Level	−1.016	.088	−11.49

tween the experts and novices in the number of target sentences selected as important. The negative estimate indicates that the novices selected more target sentences than the experts. However, the parameter estimates for the two groups differed by less than two standard errors.

The type × group interaction estimates indicate that novices were more likely to select target sentences as important when they were in the form of a definition rather than a fact. Relative to novices, the experts were more likely to select the sentences when they were in the form of a fact. These findings are shown by the negative estimate for definitions and the positive estimate for facts.

The very large parameter estimates for the main effect of level indicates that this variable had a strong effect on which target sentences were selected as important. The negative estimate shows that target sentences from the upper levels of the passage hierarchy (level 1) were selected more often than the target sentences from the lower levels (level 3). This finding is also consistent with the strong levels effect found in the ratings data. As with the ratings data, level did not interact with sentence type or group in influencing sentence selection.

Sentence selection data: Control groups. The logistic regression for the control group data indicated that the best-fitting model was a hierarchical model including the main effect of group and the type × level interaction ($\chi^2 = 1.78$, $df = 5$, $p < .88$). The inclusion of the type × group interaction reduced the fit of the model to the data. The ANOVA also indicated no type × group interaction in the data ($F(1,46) = .13$, $p < .72$) (see means in Table 2). Thus there was no evidence that the importance judgements of the undergraduate and graduate control groups differed in the degree to which they were influenced by sentence type. This can be seen in Fig. 1d, which shows the predicted mean proportion of definitions and facts selected by the two control groups plotted on a logit scale.

These means are based on the model with the type × group interaction included.

The parameter estimates for the regression model including the effect of group and the type × level interaction are presented in Table 6. Unlike the experimental groups, level did not have a linear effect on the number of target sentences selected by the control groups. It was therefore entered into the analysis as a categorical (as opposed to a linear) variable, and separate parameter estimates were obtained for each level. The parameter estimates presented for the main effect of level represent the size of the difference between levels 1 and 2, and levels 2 and 3. For the type × level interaction, the parameter estimates indicate the size of the difference between definitions and facts at each level.

The negative parameter estimate for the main effect of type shows that the control groups were more likely to select a target sentence when it was in the form of a definition than a fact. However, the type × level interaction indicates that the effect of type varied with level. The negative parameter estimate for level 1 indicates that facts were selected more often than definitions at the top level, while the positive parameters for levels 2 and 3 show that definitions were more likely to be selected than facts at the lower levels. The predicted cell

TABLE 6
PARAMETER ESTIMATES. STANDARD ERRORS. AND RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A LOGISTIC REGRESSION MODEL OF THE SENTENCE SELECTION DATA FROM THE CONTROL GROUPS (EXPERIMENT I)

Effect	Coeff.	St. Error	Ratio: Coeff. St. Error
Type	−.220	.098	−2.24
Type × Level			
Level 1	−.448	.139	−3.22
Level 2	.014	.139	.10
Level 3	.462	.140	3.30
Group	−.590	.098	−6.02
Level			
Level 1–2 diff	.476	.139	3.42
Level 2–3 diff	1.316	.140	9.40

means for this interaction are presented in Fig. 1e. Figure 1e shows that the main effect of sentence type is due primarily to a very large sentence form effect at level 3. This finding suggests that subjects without physics training may tend to judge details (low-level information) in physics texts as being more important when they are presented as definitions than facts. However, this sentence type difference is based on very few data points, as most level 3 sentences were not selected as important. The mean number of sentences from level 3 selected by the control groups were 2.3 for the undergraduates and 1.6 for the graduates. Additionally, this type × level interaction was not found in the ratings data for the control groups. Therefore, it is possible that this particular effect is not replicable.

The negative parameter estimate for the main effect of group indicates that undergraduates selected more target sentences as important than the graduate control group. This is consistent with the finding in the ratings data that undergraduates tended to rate the target sentences higher in importance relative to the graduate controls.

The positive parameter estimates for the main effect of level indicate that averaged across sentence type, the number of sentences selected as important decreased with level. The size of the estimates show that the drop in the number of target sentences selected as important was much greater between levels 2 and 3 than between levels 1 and 2. This can also be seen in Fig. 1e.

Sentence selection data: Summary. The results of the sentence selection task analyses are consistent with the findings from the ratings data. Sentence category influenced the importance judgements of novices, with novices selecting more target sentences when they were presented in the form of a definition (see Fig. 1c). Sentence category had very little effect on the sentences selected by experts. There again was no type × group interaction in the control group data, as shown in Fig. 1d, indicating that expert–novice differences are

not due to differences in educational level. Sentence category did have some influence on the judged importance of sentences for the control groups, in that they tended to select more of the definition versions than fact versions from level 3. Thus subjects who do not have scientific backgrounds may be biased towards considering low-level definitions as important.

EXPERIMENT 2

Experiment 1 demonstrated that novice importance rules are based on category membership, and not content differences between categories. Novices specifically develop a rule that definitions are particularly important; as a result, they judge the same attributive information as more important when it is presented in the form of a definition. The purpose of the second experiment was to replicate and extend these findings by examining the influence of sentence form on the judged importance of quantitative relations.

Quantitative relations are particularly important for understanding physics in that they are central to solving problems. The ability to recognize problem-relevant quantitative information is essential for efficient problem solving. However, quantitative relations vary in the manner in which they are presented in texts. They can be presented as equations, a form which explicitly signals the quantitative nature of the content, or they can be written out as verbal formulae. For example, the quantitative relation between a, b, and c, can be expressed as "$a = b/c$" (an equation) or written out as "a is equal to b divided by c" (a verbal statement).

Efficient and accurate problem solving requires that the student identify those quantitative relations relevant to the problem, regardless of the form in which they are presented. However, the results of research with uncontrolled passages (Dee-Lucas & Larkin, 1986) suggest that novices may consider equations, like definitions, to be a particularly important category of information. That is, novices may have a rule

that equations are particularly important (similar to their rule for definitions) which causes them to consider content presented in equational form as "automatically" important. If so. this importance rule could have important implications for novice problem solving performance. Novices could be biased in their assessment of the relevance of quantitative relations for problem solving depending on the form in which that information was presented.

Experiment 2 explored whether novices develop a category-based rule that equations are particularly important in physics texts. This was done by comparing expert and novice judgements of the importance of quantitative relations when they were expressed as equations or written out in verbal form. The purpose of the experiment was twofold. First. it provided a replication of Experiment 1 in support of the general finding that novices are influenced by sentence form in judging importance. Second, it extended this finding to a type of content which is central for problem solving as well as text learning in physics.

In the current study. experts and novices read passages containing target sentences which presented quantitative relations as either equations or verbal statements. Subjects rated the importance of each sentence on a 5 point scale. and selected the 10 most important sentences from each passage. The judged importance of the quantitative relations expressed as equations and verbal statements was compared to see if novices were influenced in their judgements by sentence form.

Method

Stimulus materials. The two passages were about fluid statics. and work and energy. Each was approximately 50 sentences long and each contained 9 target sentences. Examples of the equation and verbal forms of the target sentences are shown in Table 1b. Symbols for quantities were used in both versions, but the relation itself was expressed symbolically in the equation form and written out in the verbal form.

Thus the informational content was identical. and only the form in which the relation was conveyed (i.e.. symbolic or verbal) differed. There were two versions of each passage. In one version. the odd-numbered target sentences were equations and the even-numbered were verbal statements: in the other version this was reversed.

As in Experiment 1. each of the target sentences was classified according to its level in the passage hierarchy. There were 6 sentences at each of the first 3 levels in the passage hierarchies. Level was included as a variable in the data analyses to determine if perceived importance was influenced by level. and if level interacted with sentence form (i.e., equation or verbal statement).

Subjects. The novices were 18 undergraduates who had completed 2 semesters of college physics. The 18 experts had completed at least one year of graduate study in physics.

Two control groups were again included to determine if expert–novice differences were due to differences in educational level rather than differences in physics knowledge. The undergraduate control group consisted of 18 undergraduates: the graduate control group consisted of 18 graduate students who had completed at least one year of graduate study in the humanities or social sciences. None of the control group subjects had taken any college-level physics.

Procedure. The procedure and instructions were the same as for Experiment 1.

Results

The data were analyzed in the same manner as the data from Experiment 1.

Ratings data: Experimental groups. The multiway frequency analysis of the ratings data for the experimental groups indicated that the best-fitting hierarchical model included the type × group interaction and the main effect of level (χ^2 = 17.81, df = 30, $p < .96$). The ANOVA also indicated a significant type × group interaction
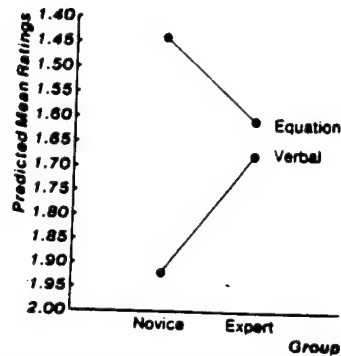
(a) Predicted mean ratings for the type x group interaction -- experimental groups.

(b) Predicted mean ratings for the type x group interaction -- control groups.
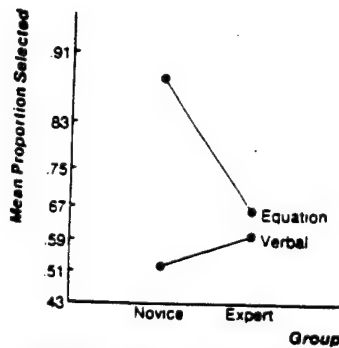
(c) Predicted proportion of target sentences selected as important for the type x group interaction -- experimental groups.
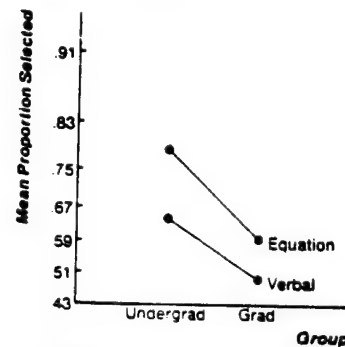
(d) Predicted proportion of target sentences selected as important for the type x group interaction -- control groups.
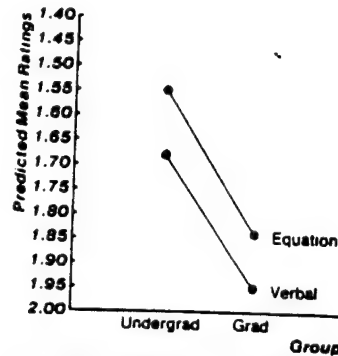
(e) Predicted proportion of target sentences selected as important for the type x level interaction -- control groups.

Fig. 2. Comparisons of the judged importance of equations and verbal formulae (Experiment 2).

$(F(1,34) = 9.17, p < .01)$. Figure 2 shows the mean ratings predicted by the model for the type × group interaction. Table 7 presents the acutal means. As in Experiment 1, novices were influenced by sentence form while experts were not. Novices considered the same information to be more important when it was expressed as an equation as opposed to a verbal statement. The predicted means for the novices were 1.44 for the equation versions and 1.92 for the verbal versions. For experts, these means were 1.61 and 1.68, respectively.

The parameter estimates for the complete model are shown in Table 8. The esti-

TABLE 7
MEAN IMPORTANCE RATINGS AND MEAN PROPORTION OF TARGET SENTENCES SELECTED AS IMPORTANT FOR THE TYPE × GROUP INTERACTIONS (EXPERIMENT 2)

	Experimental groups		Control groups	
	Novice	Expert	Undergrad	Grad
Ratings				
Equation	1.32	1.54	1.47	1.79
Verbal	1.88	1.65	1.67	1.92
Proportion selected				
Equation	.84	.65	.75	60
Verbal	.49	.64	.61	.47

TABLE 8
PARAMETER ESTIMATES. STANDARD ERRORS. AND RATIO OF ESTIMATES TO STANDARD ERRORS FOR A LOGLINEAR MODEL OF THE RATINGS DATA FROM THE EXPERIMENTAL GROUPS (EXPERIMENT 2)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
(a) Group: Novice estimates (expert estimates are opposite)			
Rate 1	-.016	.081	-.20
Rate 2	.116	.098	1.19
Rate 3	.003	.125	.02
Rate 4	-.187	.182	-1.03
Rate 5	.085	.205	.42
(b) Type: Equation estimates (verbal estimates are opposite)			
Rate 1	.297*	.083	3.60
Rate 2	-.154	.098	-1.58
Rate 3	-.031	.119	-.26
Rate 4	-.195	.183	-1.07
Rate 5	.083	.205	.40
(c) Type × Group: Equation (verbal estimates are opposite)[a]			
Rate 1 Novice (+) Expert (-)	.236*	.083	2.86
Rate 2 Novice (Expert (-)	.035	.098	.36
Rate 3 Novice (+) Expert (-)	.027	.120	.23
Rate 4 Novice (-) Expert (+)	-.256	.182	1.40
Rate 5 Novice (-) Expert (+)	-.042	.205	-.21
(d) Level			
Level 1			
Rate 1	1.065*	.164	6.50
Rate 2	-.216	.200	-1.08
Rate 3	-.233	.249	-.94
Rate 4	-.461	.417	-1.11
Rate 5	-.155	.423	-.37
Level 2			
Rate 1	-.087	.135	-.64
Rate 2	.352*	.154	2.29
Rate 3	.158	.191	.83
Rate 4	-.252	.323	.78
Rate 5	-.170	.348	-.49
Level 3			
Rate 1	-.978*	.122	-8.01
Rate 2	-.136	.139	-.98
Rate 3	.075	.172	.44
Rate 4	.713*	.264	2.70
Rate 5	.325	.291	1.12

Note. Asterisks indicate coefficients that differ from zero by more than 2 standard errors.
[a] Parameter estimates indicate the size of the effect: positive and negative symbols indicate the direction of the effect for each group.

mates for the main effect of group show little difference between experts and novices in their use of the different ratings. Novices had a slightly greater tendency to use the rating 2 category relative to experts. while experts tended to use the rating 4 category more often. but the magnitude of these effects are not large (i.e.. do not differ from zero by more than 2 standard errors).

The estimates for the main effect of type show a large difference in the overall ratings of equations and verbal formulae. Equations were much more likely than verbal statements to receive the highest importance rating (rating 1). and somewhat less likely to receive a rating of 2 and 4. This indicates that equations were rated higher in importance overall than verbal formulae.

The type × group interaction estimates indicate that novices were much more likely to give a rating of 1 to equations than verbal formulae. and somewhat more likely to give a low rating of 4 to verbal formulae than equations. The opposite was true for experts relative to novices.

The level parameter estimates again show that level has a strong effect on judged importance. Subjects tended to give level 1 target sentences a rating of 1, level 2 target sentences a rating of 2. and level 3 target sentences a rating of 4. Level did not interact with type or group.

*Ratings data: Control groups.* The multiway frequency analysis of the ratings data for the control groups indicated that the most appropriate model included the main effects of type. level, and group (χ^2 =

23.40, $df = 35$, $p < .93$). The inclusion of the type × group interaction in the model reduced its fit to the data set. The ANOVA also indicated no type × group interaction ($F(1,34) = .24$, $p < .62$) (see means in Table 7). Thus the effect of sentence form on rated importance did not vary with level of expertise. This can be seen in Fig. 2b, which shows the mean ratings predicted by the model with the type × group interaction included. The difference in the ratings of the equation and verbal sentence forms did not vary with the educational level of the control groups.

The parameter estimates for the model including the main effects of group, type, and level are shown in Table 9. The estimates for the main effect of sentence type indicates that sentence form did influence the importance ratings of the control groups. Equations were more likely to be rated 1 and verbal formulae rated 2. This can also be seen in Fig. 2b, which shows that both groups judged information as more important when it was presented in the equational form as opposed to the verbal form. This was not the case in Experiment 1, in that the judged importance of the target content did not vary when it was presented in the form of a definition or fact.

The estimates for the main effect of group show that there were large differences between the two groups in their use of ratings 1, 3, and 4. The undergraduate controls tended to rate the sentences either very high in the importance (1) or very low in importance (4), while the graduate controls were more likely to use the intermediate rating (3).

The estimates for level show the usual level effects. Subjects rated the top-level target sentences high in importance (rating 1), and low level sentences low in importance (rating 5). The ratings for level 2 sentences did not tend to cluster in any one rating category.

Ratings data: Summary. The ratings

data comparing the judged importance of equations and verbal statements confirm the findings from Experiment 1. In both experiments, novices were sensitive to sentence form while experts were not. Specifically, when sentence information was signaled as belonging to a certain information category (i.e., definition or equation),

TABLE 9
PARAMETER ESTIMATES, STANDARD ERRORS, AND RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A LOGLINEAR MODEL OF THE RATINGS DATA FROM THE CONTROL GROUPS (EXPERIMENT 2)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
(a) Group: Undergrad estimates (grad estimates are opposite)			
Rate 1	.280*	.083	3.39
Rate 2	.018	.099	.18
Rate 3	−.301*	.115	−2.63
Rate 4	−.314*	.135	−2.33
Rate 5	−.312	.240	−1.30
(b) Level			
Level 1			
Rate 1	.482*	.133	3.62
Rate 2	.036	.162	.22
Rate 3	.097	.172	.56
Rate 4	−.213	.227	−.94
Rate 5	−.402	.417	−.97
Level 2			
Rate 1	−.001	.250	−.01
Rate 2	.067	.143	.47
Rate 3	−.100	.157	−.64
Rate 4	.052	.193	.27
Rate 5	−.019	.352	−.05
Level 3			
Rate 1	−.482*	.110	−4.37
Rate 2	−.102	.133	−.77
Rate 3	.003	.158	.02
Rate 4	.161	.179	.90
Rate 5	.420	.303	1.39
(c) Type: Equation estimates (verbal estimates are opposite)			
Rate 1	.150	.078	1.92
Rate 2	−.181	.098	−1.84
Rate 3	.008	.103	.08
Rate 4	.104	.131	.79
Rate 5	−.082	.220	−.37

Note. Asterisks indicate coefficients that differ from zero by more than 2 standard errors.

novices rated that content higher in importance than when category membership was changed to that of another category. This "form effect" reflects novice rules regarding the importance of different types of content in physics texts.

Unlike Experiment 1, the control group data showed that physics-naive subjects were also influenced by sentence form in the presentation of quantitative relations. There are two possible explanations for this sentence form effect in the control groups. First, there may be general rules formulated by the population at large regarding the relative importance of different types of content in certain types of texts. Physics-naive as well as novice subjects may consider equations to be particularly important in this type of text. This suggests that there may be a general consensus that equations are particularly important in science texts. Note, however, that this initial preconception about the importance of equations is strengthened by a limited amount of physics training. The difference in the rated importance of the equation and verbal forms of the target sentences is much greater for novices than physics-naive subjects (see Figs. 2a and 2b).

It is also possible that differences in the spatial format of the equations and verbal statements influenced the importance judgements of the controls. The verbal statements were presented as continuous text, while the equations were indented and presented on a separate line. In general, text information that is made distinctive is better recalled, presumably because it is assumed to be important and therefore receives more attention (Cashen & Leicht, 1970; Fowler & Barker, 1974; Glynn & DiVesta, 1979; Lorch & Chen, 1986). Thus the controls may have judged equation versions as more important because they were offset in the text. In this case, the form effect for control subjects would reflect a general reading strategy of assuming that spatially prominent text information was particularly important.

This raises the possibility that the novices' importance judgements were also based on differences in typographical format, rather than a rule that equations are more important than facts. However, as shown in Figs. 2a and 2b, the difference in the judged importance of equational and verbal statements was much larger with novices than controls. If this difference was due to general reading strategy alone, then there is no reason to expect the difference to be greater for people who are familiar with physics (novices) than people who are unfamiliar (controls). The larger form effect for novices indicates that this effect is partially knowledge-dependent. It suggests that novices specifically develop a rule that equations are particularly important, and consistently judge equations accordingly.

The expert data show that after extensive physics training, subjects recognize the importance of quantitative relations regardless of whether they are presented as equations or verbal formulae, and thus are not influenced by sentence form. The lack of a type × group interaction in the control group data again indicates that expert-novice differences are not due to differences in educational level.

Sentence selection data: Experimental groups. The logistic regression for the sentence selection data from the experimental groups indicated that the best-fitting model was a hierarchical model including the type × group interaction and the main effect of level ($\chi^2 = 10.32$, $df = 7$, $p < .17$). This is the same model that accounted for the experimental group data from Experiment 1 and the experimental ratings data in the current study. The ANOVA also indicated a significant type × group interaction ($F(1,34) = 16.74$, $p < .01$). The predicted mean proportion of target sentences selected for each group and type are shown in Fig. 2c plotted on a logit scale. The actual means are in Table 7. This interaction shows that novices are much more likely to select a target sentence as important when

it is presented in the form of an equation as opposed to a verbal formula, while experts' sentence selection is not influenced by sentence category. This is consistent with the results of Experiment 1 and the ratings data indicating that novices are influenced by sentence form, while experts are not.

The parameter estimates for the logistic regression model are shown in Table 10. As in Experiment 1, these indicate the size of the difference between the means for an effect. The negative parameters for the main effects of type and group indicate that summed across groups, equations were selected more often than verbal formulae, and that novices selected more of the target sentences overall than the experts.

The type × group interaction estimates show that novices were more likely to select target sentences as important when they were presented as equations as opposed to verbal formulae. Relative to novices, experts were more likely to pick sentences in the verbal form, as indicated by the positive parameter for verbal formulae and the negative parameter for equations.

The large negative parameter estimate for level replicates previous results indicating that subjects are much more likely to select high-level sentences than low-level sentences as important.

Sentence selection data: Control groups.

The most appropriate model for the logistic regression for the control group data was a hierarchical model including the main effect of group and the type × level interaction ($\chi^2 = 7.02$, $df = 5$, $p < .22$). The inclusion of the type × group interaction resulted in a poorer fit of the model to the data. The ANOVA also indicated no type × group interaction ($F(1.34) = .01$, $p < .91$) (see means in Table 7). Thus the two control groups did not differ from each other in their evaluation of the importance of the equations and their verbal equivalents. This can be seen in Fig. 2d, which shows the predicted mean proportion of target sentences selected for the type × group interaction, based on the model with the interaction term included. Figure 2d shows that both undergraduates and graduate students picked equations more often than verbal formulae. This is consistent with the results of the ratings data (shown in Fig. 2b).

The parameter estimates for the regression model including the main effects of group and type × level interaction are shown in Table 11. As in Experiment 1, level did not have a linear effect on the number of target sentences selected by the control groups, and therefore was entered as a categorical variable in the analysis. The parameter estimates show the size of the difference between levels. The type × level parameter estimates indicate the size of the difference between equations and verbal formulae for each level.

The parameter estimate for the main effect of type shows that overall the controls selected more equations than verbal formulae as important. However, the type × level interaction indicates that the effect of sentence form varied with level, with the largest difference between sentence types occurring at level 2. Figure 2e shows the predicted mean proportion of sentences selected for this interaction. The preference for equations n be seen at levels 1 and 2, with little difference in sentence type selection at level 3. Thus the general tendency

TABLE 10
PARAMETER ESTIMATES. STANDARD ERRORS, AND RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A LOGISTIC REGRESSION MODEL OF THE SENTENCE SELECTION DATA FROM THE EXPERIMENTAL GROUPS (EXPERIMENT 2)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
Type	−1.100	.137	−8.03
Group	−.530	.134	−3.96
Type × Group Verbal (+) Equation (−)	.842	.136	6.19
Level	−1.239	.126	−9.82

TABLE II

PARAMETER ESTIMATES. STANDARD ERRORS. AND
RATIOS OF ESTIMATES TO STANDARD ERRORS FOR A
LOGISTIC REGRESSION MODEL OF THE SENTENCE
SELECTION DATA FROM THE CONTROL GROUPS
(EXPERIMENT 2)

Effect	Coeff.	St. Error	Ratio: Coeff./ St. Error
Type	-.544	.120	-4.52
Type × Level			
Level 1	.262	.176	1.49
Level 2	-.652	.171	-3.82
Level 3	.390	.165	2.36
Group	-.756	.120	-6.28
Level			
Level 1-2 diff	.611	.174	3.51
Level 2-3 diff	.530	.168	3.15

for physics-naive subjects to consider equations as important may occur primarily for higher-level text information rather than text details. However, this interaction was not found in the ratings data and may in part reflect constraints imposed by the limited choice (i.e., pick 10) dependent measure. While many sentences were chosen from levels 1 and 2, relatively few sentences of any type were selected from level 3. Thus the lack of a difference at level 3 may be partly due to a floor effect.

The negative parameter for the main effect of group shows that the undergraduates tended to select more target sentences as important than the graduate students. This is consistent with the ratings data indicating that the undergraduates rated the target sentences more important overall than the graduate controls.

The two positive parameter estimates for level indicate that the controls selected more sentences from the higher levels than lower levels of the texts. The size of the coefficients show that summed across sentence types, the decrease between levels 1 and 2 in the number selected was roughly equal to the decrease between levels 2 and 3. This can also be seen in Fig. 2e.

Sentence selection data: Summary. The

data from the sentence selection task again support the findings from Experiment 1 that novices are sensitive to variations in sentence form when these variations signal specific category membership. Novices judged the same quantitative relations as more important when they were presented as equations as opposed to being written out in verbal form. Experts were not influenced in their importance judgements by sentence form.

As in the ratings data. the control subjects also considered the equational form of the target sentences to be more important. This effect was confined to the first two levels of the passage hierarchy in the sentence selection data. This again could indicate that physics-naive people also consider equations to be particularly important in science texts. or reflect the fact that readers tend to judge as important content emphasized by the spatial layout of the text. As with the ratings data. the influence of sentence form was much greater for novices than physics-naive subjects. indicating that this effect for novices is not solely based on differences in spatial prominence.

DISCUSSION

The research reported here examined one aspect of novice knowledge representations—rules for assessing importance in unfamiliar scientific domains. The results confirm that novice importance rules identified in earlier research using uncontrolled materials are based on category membership. and not content differences between categories. Novices consider the same substantive information to be more important when presented as a definition (rather than a fact) and as an equation (rather than a verbal phrase). Thus novices are sensitive to variations in sentence form when these variations indicate membership in particular information categories. This "form effect" is not seen in experts, and is either absent or attenuated in physics-naive sub-

jects. Therefore the apparent form effect seen in experts with uncontrolled materials reflect differences in content, not form— when the content of different categories is held constant, experts judge the same information presented in different forms to be equivalent in terms of its importance. This indicates that experts judge certain categories as more important than others because these categories typically contain the type of content that is important for understanding physics. Thus novices judge importance by the information category, while experts judge according to content importance, which is correlated with category in natural text.

The naive control subjects in these studies were also relatively uninfluenced by sentence category, and therefore appeared similar to experts. This is most likely because they have few expectations about what types of content are important in physics. Thus experts are not influenced by sentence category because their highly refined knowledge allows them to assess importance on the basis of content; naive subjects are not influenced by category because they lack the basic knowledge necessary to develop hypotheses as to what is important in the domain.

These findings indicate that people just beginning to learn about a content domain develop rules specifying what categories of information are important in that domain. This suggests how a content schema might evolve in novice learners. An early stage in content schema development may be the specification of rules indicating what information categories are important in the domain. With increasing expertise, novices could begin to differentiate the important and less important information within categories, and thus move from a classification system based on information categories to a more differentiated expert knowledge structure based on a deeper analysis of the nature of the category content. This type of schema shift has been found in novice classifications of math and physics problems (Chi. Feltovich. & Glaser. 1981; Schoenfeld & Herrmann. 1982). Novices classify problems on the basis of superficial characteristics of the problems; experts classify according to underlying principles used in problem solving; and subjects of intermediate expertise use both dimensions in their classification systems (Chi et al.. 1981). The current research suggests a similar shift in the text features viewed as relevant for assessing importance. Novices consider the surface-level feature of category membership as pertinent to judging importance; experts rely on a deeper analysis of the nature of the text content.

The category-based rules developed by novices have important consequences for what novices learn from texts. Previous research has found that novices spend more time on information categories judged as important when reading physics passages, recall more information from categories judged as important, and include more information from these categories in their summaries of physics texts (Dee-Lucas & Larkin. 1986. in press). Thus these information-category rules appear to influence novice readers' attentional processes during reading, as well as the macrostructure they develop for texts. The specific findings from the current research suggest that novices may be missing important facts and quantitative relations, and attending to some less important definitions and equations when studying these types of texts.

Novice rules may also have implications for novice problem-solving performance. The finding that novices consider equations to be more important than verbal formulae suggests that novices may not readily recognize quantitative relations as being relevant in a problem-solving situation when this information is not presented in equational form. Similarly, novices may be more likely to attend to irrelevant quantitative relations when they are presented as

equations. This is supported by problem-solving research indicating that novices tend to invoke equations early in the problem-solving process. suggesting that novices are in fact "drawn" to this type of content (Larkin. 1981. 1983). The inability to pinpoint problem-relevant information would decrease the efficiency (and possibly accuracy) with which novices solve physics problems.

REFERENCES

CASHEN. V. M.. & LEICHT. K. L. (1970). Role of the isolation effect in a formal educational setting. *Journal of Educational Psychology.* 61, 484–486.

CHI. M.. FELTOVICH. P.. & GLASER. R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science. 5,* 121–152.

DEE-LUCAS. D.. & LARKIN. J. H. (1986). Novice strategies for comprehending scientific texts. *Discourse Processes.* 9, 329–354.

DEE-LUCAS. D.. & LARKIN. J. H. (in press). Attentional strategies for studying scientific texts. *Memory & Cognition.*

FIENBERG. S. E. (1980). *The analysis of cross-classified categorical data.* Cambridge. MA. MIT Press.

FOWLER. R. L.. & BARKER. A. S. (1974). Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology.* 59, 358–364.

GLYNN. S. M.. & DiVESTA. F. J. (1979). Control of prose processing via instructional and typographical cues. *Journal of Educational Psychology.* 71, 595–603.

KIERAS. D. E. (1985). Thematic processes in the comprehension of technical prose. In B. K. Britton and J. B. Black (Eds.). *Understanding expository text* (pp. 89–107). Hillsdale. NJ: Erlbaum.

LARKIN. J. H. (1981). Enriching formal knowledge: A model for learning to solve problems in physics. In J. R. Anderson (Ed.). *Cognitive skills and their acquisition* (pp. 311–334). Hillsdale. NJ: Erlbaum.

LARKIN. J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.). *Mental models* (pp. 75–98). Hillsdale. NJ: Erlbaum.

LORCH. R. F.. & CHEN. A. H. (1986). Effects of number signals on reading and recall. *Journal of Educational Psychology.* 78, 263–270.

SCHOENFELD. A. H.. & HERRMANN. D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology.* 8, 484–494.

Appendix C

(Total pages in length: 40)

Dee-Lucas, D. & Larkin, J.H. (1989a)

Organization and comprehensibility in scientific proofs, or "Consider a particle p...".

(Tech. Rep. A.C.P. 19). Carnegie Mellon University

Organization and Comprehensibility in Scientific Proofs,

or "Consider a particle p ..."[*]

Diana Dee-Lucas & Jill H. Larkin

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

A.C.P. #19

May 1989

Running head: Organization and Comprehensibility

Abstract

Technical texts often present scientific principles by first presenting a proof of the principle, and only stating the principle after the proof—a "proof-first" organization. This specialized text structure differs from conventional structures in that it does not provide readers with thematic information to guide their processing of the text. The current research examined the effects on comprehension of this proof-first organization. This was done by comparing the processing of proof-first texts to that of texts having the more conventional "principle-first" structure, in which the passage theme (i.e., the principle) is stated at the beginning of the text. Readers found the principle-first structure easier to process, and used it to summarize texts presented in both principle-first and proof-first structures. Additionally, the principle-first structure enhanced immediate and delayed recall of the principle. These results are discussed in terms of differences in the processing requirements of these two structures, and their implications for writers.

Organization and Comprehensibility in Scientific Proofs, or "Consider a particle p ..."

Theories of text processing suggest that reading comprehension is facilitated by conventional rhetorical structures. Readers can use their knowledge of these typical structures to organize and anticipate the text content during reading (Mandler, 1978; Meyer, 1983; van Dijk & Kintsch, 1983). Some common rhetorical structures include argument, comparison and contrast, problem and solution, classification, procedural description, etc. Structures such as these are common to writing in many different content areas, and hence are familiar to readers of varying backgrounds. However, there are also specialized text structures used in organizing information in particular content domains. These include, for example, the structure of legal documents (Danet, 1980) and scientific reports (Vesonder, 1979). These structures are limited to certain topic areas, and hence generally unfamiliar to most readers.

Specialized structures are first encountered by readers when first learning about a new domain. Because of their unfamiliarity, these structures may make reading more difficult. A familiar structure provides a known organizational framework within which to incorporate the text content. Once readers recognize the form in which the content is being presented, they can simply match the content to that organizational schema as they read. In contrast, when readers are not familiar with the text structure, they must determine the text structure at the same time that they are trying to understand the content itself. This additional processing may slow comprehension and interfere with learning.

The current research investigated the effects on comprehension of a specialized text structure found primarily in science and mathematics texts. This organization is used to present principles and their associated proofs. In presenting this type of content, it is common to present the proof prior to the statement of the principle, a "proof-first" structure. That is, the text will present a hypothetical situation and proceed to derive a principle or rule using the elements in that situation, with no initial statement of the principle to be derived. The proof thus serves as the introduction of the principle— the reader does not know what is being proven until reaching the end of the proof. The logical alternative to this common proof-first format is an organization in which the principle is presented prior to its proof, a "principle-first" organization. With this organization, the principle is stated and then proven, so the reader knows in advance the goal of the proof. The research reported here compared the comprehension of scientific texts having these two organizations by readers who are unfamiliar with scientific materials (i.e., novices). Although the only difference between these two structures is in information order, there are several reasons to expect that novice readers would find the scientific proof-first structure more difficult to process.

First, the proof-first structure fails to signal directly the passage organization. There are typically no explicit rhetorical cues indicating that the text consists of a proof followed by a principle—the text simply begins with the proof itself (usually a description of the general setting, as in "Consider a particle p ..."). Novices, who by definition have not frequently encountered this type of structure, would not possess a text schema reflecting the structure of a proof-first text. That is, they would have no way to recognize this type of introduction as a signal for a proof-then-principle organization, and therefore would be unable to anticipate the overall organization of the text.

Second, the proof-first structure does not provide novice readers with an initial means of determining the main ideas of the passage. Without knowing the principle, the reader does not know the point of the passage, or the direction that the proof will take. Several theories of prose comprehension have emphasized the importance of having available early in a text, thematic information that can be used as a superordinate conceptual organizer for processing more detailed text content (Ausubel, 1963; Bransford & Johnson, 1972; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). For example, in the Kintsch and van Dijk text processing model, readers beginning a text select an initial proposition to use as a superordinate for processing subsequent content. All subsequent propositions are processed in terms of their relationship to this initial proposition. In the principle-first structure, the statement of the principle can be used as a superordinate concept to which the proof is related in a subordinate manner. However, the proof-first structure does not provide the reader with this initial superordinate content. Although the reader knows the general topic, the main point of the passage (the principle) is not presented until the end. Without knowing the principle in advance, readers would have more difficulty organizing the information around a central idea as they are reading, and in particular may have more difficulty determining what is important (Kintsch & Yarborough, 1982).

The scientific proof-first structure would be expected to be difficult for novices to process because it lacks structural and thematic cues. Additionally, the proof-first structure violates readers' expectations as to how information should be ordered in a text. Readers expect important information to occur at the beginning of a text (Kieras, 1979), and in the proof-first structure the most important content (the principle) occurs last. The end of the text, however, is the second most likely location in which readers expect important content to occur (Kieras, 1979). This observation provides some empirical evidence that structures consisting of an argument leading up to a conclusion, such as the proof-first structure, are not completely foreign to readers, even though they are less common (Kieras, 1980b).

The three studies reported here contrasted the principle-first and proof-first organizations in terms of their ease of comprehension for novice readers. This was done by constructing several pairs of passages containing principle and proof information presented in both organizations. To conform

with other norms of scientific text, both versions began with an informative title and paragraphs introducing the general topic. Thus the passages were not like those used in research examining the effects of scrambled passages (Kintsch, Mandel, & Kozminsky, 1977; Kintsch & Yarbrough, 1982), passages without topic sentences or clearly defined topics (Kieras, 1980b), or passages lacking titles defining an ambiguous passage subject (Bransford & Johnson, 1973). Those studies used unnatural texts in order to demonstrate specific text processing principles. In contrast, the purpose of the current research was to compare the effects on text processing of two expository text organizations that differ in information order and in familiarity of text structure. All passages clearly indicated the subject matter and topic being discussed. The difference was whether or not knowledge of the principle was available before reading the corresponding proof.

Three studies examined differences produced by these alternative text structures in novice readers' attention, in their judgements of what is important, and in the organization of content in summaries and passage recall. The first experiment examined effects on reading times and on the organization of readers' text summaries. The second experiment examined differences in the perceived importance of the information, and in the ability of readers to predict what is important as they are reading. The third experiment investigated differences in immediate and long-term recall of text content.

Experiment 1

The purpose of the first experiment was to determine whether the two structures produce attentional differences, and if novices use the more familiar principle-first structure preferentially in their own summaries. Subjects having little knowledge of physics read physics texts presented in these two information orders while their sentence reading times were recorded. Afterwards they wrote brief summaries of the passages. Reading times were examined to see if readers found one organization easier to process overall, and if they attended to different types of information in the two structures. The summaries were analyzed to determine whether readers tended to maintain the original passage organization, or reorganize the text according to a preferred structure.

Method

Materials. Four target passages were constructed, each containing a principle and a proof segment. The passages were written so that the two main segments were interchangeable; that is, the proof or the principle could be presented first. The passages ranged from 19 to 24 sentences in length. Each passage began with an introductory buffer paragraph (4 or 5 sentences in length) introducing the passage topic, followed by the principle and proof in the appropriate order, and ending with a closing paragraph of 3 or 4 sentences.

The principle segment contained a statement of the main principle along with introductory and closing statements, for a total of 3 or 4 sentences. The proof segment was necessarily longer

because it contained all the steps of the proof argument. The proof began with a few introductory statements describing the general setting (e.g., "Assume that there is a liquid in equilibrium in a closed cylinder that is fitted with a piston..."). These were followed by the proof itself, ending with a statement of the principle (i.e., the result of the proof). This statement of the principle within the body of the proof was considered to be the "core" sentence of the proof in that it was the goal of the proof. The proof segment was 10 or 11 sentences in length. Two of the passages included diagrams with some of the proof sentences. Examples of the two versions of one of the passages are shown in Table 1.

```
┌─────────────────────────────────────────────────┐
│              Insert Table 1 about here           │
└─────────────────────────────────────────────────┘
```

Subjects. The subjects were 24 undergraduates who had completed no more than 1 semester of college-level physics.

Procedure. The study was conducted in two sessions. In each session subjects read two practice passages and two target passages. The order of presentation of the practice passages was kept constant and the order of the target passages was counterbalanced. Each session alternated practice and target passages, beginning with a practice passage. The passages were presented on a VDT one sentence at a time. Subjects pressed a button to display the next sentence.

Subjects were told that they would be reading short passages about physics and then writing an 8 to 12 sentence summary of each. The summaries were to contain what they considered to be the main points of each passage. After completing the reading of each passage (and before writing their summary), subjects were asked to rate their prior familiarity with the passage topic on a scale from 1 to 5 (1 being very familiar). At the end of each session, subjects rated the individual sentences in the target passages according to how familiar they had been with the content prior to reading the passage. These ratings were to have been used in the analysis to control for within-subject variation in prior knowledge of the passage information. However, the inclusion of the ratings in the reading time analyses did not alter the results, so these ratings were not used.

Results

The mean familiarity ratings for the four passages indicated that one of the passages was much more familiar to subjects than the others. The mean rating for this passage was 2.58 on the 1 to 5 scale (1 being very familiar). Preliminary analyses of the reading time and summary data indicated that this passage produced a data pattern qualitatively different from those of the other passages, which had lower familiarity ratings. Therefore, the data for this passage were excluded and all analyses were conducted on the data from the remaining three passages.

Two of the target passages contained diagrams which were referred to in a total of three sentences. Because the diagram inspection time was confounded with sentence reading time and could increase the variability in reading times, the data for these sentences were dropped from the analysis. With these omissions, the final data set contained 960 reading times.

Reading times: Principle and proof. The reading times were analyzed with a multiple regression performed on the logarithm of the sentence reading times expressed in milliseconds. The variables included were experimental session (first/second), passage structure (principle-first/proof-first), sentence type (principle/proof), number of words, passage, and the type x structure interaction. This interaction reflects the hypothesized differential processing of principle and proof sentences depending on the order in which they are read. Because the analysis was performed on the individual reading times for each subject on each sentence, subjects were also included as a categorical variable. The regression coefficients for this model, excluding the subject variables, are presented in Table 2. Results indicated that all effects were significant except the main effect of structure. Figure 1a summarizes the type x structure interaction. It shows (on a log scale) the reading times predicted by the statistical model for sentences of average length in the passages (i.e., 24 words) and averaged over session and passage. Figure 1a shows that readers spend more time on the same text content when it is presented at the beginning of the passage. The principle-first organization results in more time being spent on principle information; the proof-first structure produces longer reading times for the proof. This suggests that information is more thoroughly processed when it is placed at the beginning of a text.

Insert Table 2 and Figure 1 about here

The coefficient for the main effect of structure shows that readers spent about the same amount of time on the sentences (averaged across principle and proof) in both structures. The coefficient for type indicates that readers spent more time overall on the sentences in the proof segment as opposed to the principle. This result most likely reflects differences in syntactic and/or semantic characteristics of the principle and proof information. The passage coefficients indicate that the overall reading time varied with the passage topic, with some passages being read faster than others. The positive session coefficient shows that readers read the passages more slowly during the first experimental session. Finally, the coefficient for words indicates that readers spent more time on longer sentences.

Reading times: Different information types. The analysis described above compared reading times for the two component text segments of interest, the principle and proof. In order to determine whether text structure influenced the reading times for subsets of information within these

larger segments, the text content was divided into four types: core principle sentences, core proof sentences, introductory information, and non-core proof sentences. These were defined in the following way.

"Core principle" sentences were located in the principle segment and expressed the principle derived in the passage. There were one or two core principle sentences in each passage.

"Core proof" sentences were the goal statements of the proofs, equivalent in content to the core principle sentences, but expressed in different words. There was one core proof sentence per passage. In principle-first passages, the core proof was a restatement of the core principle statement occurring earlier in the passage; in proof-first passages, the core proof sentence was the initial statement of the principle.

"Introductory" sentences contained non-substantive passage content. They either (a) introduced the principle (e.g., "Pascal's principle was discovered in the seventeenth century by the French philosopher, mathematician, and physicist, Blaise Pascal"), (b) served as a transition between the principle and proof (e.g., "This is a necessary consequence of the laws of fluid mechanics, as shown below"), or (c) provided background information for the proof (e.g., "Consider an arbitrary portion of a fluid at rest.").

"Proof" sentences were the component sentences of the proof leading up to the core proof statement (e.g., ". . . Because the fluid portion is at rest, the upward buoyant force must exactly balance the downward gravitational force. Therefore the buoyant force must be directed vertically upward and be equal in magnitude to the weight of the fluid in V ..."). These consisted of all sentences in the proof segment which were neither introductory sentences nor the core proof sentence.

Each principle segment consisted of the core principle sentences and introductory sentences. The proof segment consisted of introductory sentences, proof sentences, and the core proof statement. The total number of sentences falling into these four categories (summed across passages) were: core principle: 5; core proof: 4; proof: 18; and introductory: 13.

A multiple regression analysis was performed on the logarithm of the reading times for sentences falling into these four categories. As before, the three sentences referring to figures were excluded from the analysis. The variables included sentence type (core principle/core proof/proof/introductory), structure (principle-first/proof-first), passage, session, number of words, and the type x structure interaction for each of the four sentence types. Subjects were also included as a categorical variable. The coefficients from the analysis (excluding those for subjects) are presented in Table 3. The analysis shows that passage structure had the greatest effect on reading times for the core principle sentences. Figure 1b summarizes the type x structure interactions. It shows (on a log scale) the reading times predicted by statistical model for sentences of average length (i.e., 24 words) and

averaged over session and passage. The data for the two core sentences are grouped together for comparison with the interaction for the complete proof and principle segments shown in Figure 1a. Figure 1b shows that readers spend more time on the core principle content when it occurs first. This is also true for the core proof content, but the effect is smaller and less reliable. There was a consistent effect for the non-core proof sentences, but in this case the effect was smaller than for both core sentences, and also less reliable than for core principle sentences. There is no difference in the amount of time spent on introductory content in the two passage structures, indicating that readers are not influenced by information order when the content is not substantive text information.

Insert Table 3 about here

<u>Summaries.</u> The 72 summaries (3 passages, 24 subjects) were scored according to whether subjects included information from both the principle and the proof in their summaries, and if so, whether they maintained the original passage structure or reversed the proof-principle order. They were scored independently by two scorers to assess inter-scorer reliability. There was 94% agreement on the passage structure scoring. A total of 12 summaries (6 for each structure) excluded either the principle or the proof, and thus could not be analyzed for structural change.

The results indicated that when subjects included both principle and proof information in their passage summaries, there was a much greater tendency to reverse the order of the content when it had been presented in a proof-first structure. With the principle-first structure, 29 out of 30 summaries maintained the original passage structure. With the proof-first structure, only 5 of 30 summaries maintained the proof-first structure (25 summaries reversed the principle-proof order). A chi-square on these frequency counts was highly significant ($\chi^2=39.10$, $df=1$, $p<.001$). This suggests that readers find the principle-first structure more natural in that they overwhelmingly use this structure, even when summarizing proof-first texts.

<div align="center"><u>Experiment 2</u></div>

The results of the first experiment indicate that the principle-first and proof-first structures result in differential processing of the principle and proof information. The goal of the second experiment was to determine if these processing differences were due in part to differences in the perceived importance of the content. This would be consistent with previous research showing that readers expect important content to occur at the beginning of a text (Kieras, 1979).

In the current study, novices rated the importance of sentences in experimental texts as they were reading the passages (rating each sentence before reading the next sentence). After completing the entire passage, readers were allowed to go back through the passage and revise their ratings. The purpose was to determine how text structure influenced (a) the perceived importance of

sentences (as indicated by the initial importance ratings), and (b) the ability of readers to pick out the important content as they read (as indicated by the number of rating revisions).

Method

Materials. Two of the four passages used in Experiment 1 were used in Experiment 2. Some minor revisions were made in these two passages to reduce the overlap between the core principle and core proof sentences. Each passage contained one core principle and one core proof sentence. The passages were printed one line on a page so the subjects could rate the importance of each sentence prior to viewing the information following that sentence.

Subjects. The subjects were 20 undergraduates who had completed no more than two semesters of college-level physics.

Procedure. The subjects were told that this study examined how people go about determining what is important as they read. They were informed that several passages would be presented, one sentence at a time, and that they were to rate the importance of each sentence as it was presented using a 5-point scale, (1 being most important (i.e., a main point of the passage) and 5 being least important (i.e., information unrelated to the main points of the passage)). They completed the ratings for four passages, two practice and two target, with one practice passage preceding each target passage. The practice passages differed greatly from the target passages in both content and organization so as to prevent subjects from noticing the structural manipulation in the target passages. After they had completed ratings for the four passages, subjects were told that they should go back through the two target passages and revise any ratings that they felt were inaccurate based on their knowledge of the complete passage.

Results

Although subjects rated all sentences, sentences referring to diagrams were excluded from all data analyses (as was done in Experiment 1). This was done to eliminate potential variation in judging the importance of diagrams. One passage contained two sentences with diagrams and the other contained one.

Principle and proof: Mean importance ratings. The ratings for the principle and proof segments were analyzed with a multiway frequency analysis with type (proof/principle), passage, structure (proof-first/principle-first), and rating (original/revised) as factors. This analysis fits a loglinear model to categorical data. The loglinear model expresses the logarithm of the expected cell frequencies as an additive function of main effects and interactions, in a manner similar to the typical analysis of variance model (Fienberg, 1980). The analysis indicated significant type x structure and structure x passage interactions. Because there was no significant difference between the original and revised ratings, the analysis was rerun on just the original ratings. This was done in order to look at the

effect of information order on perceived importance during the reading of the passage, before subjects had full knowledge of the passage content. Results again indicated significant type x structure and structure x passage interactions (χ^2=5.52, df=10, p<.85). The coefficients for this model are presented in Table 4. The finding of interest, the type x structure interaction, is summarized by Figure 2a, which shows the mean ratings given to the principle and proof content presented in a principle-first and a proof-first structure. These means for the principle content were 2.83 with the principle-first structure and 3.02 with the proof-first structure. The corresponding means for the proof were 2.64 (principle-first) and 2.54 (proof-first). Readers rated the same content higher in importance when it was presented at the beginning of the passage. This is consistent with research indicating that readers tend to expect the important information to occur early in a text (Kieras, 1979). The structure x passage interaction was due to the fact that subjects rated sentences in the principle-first version higher in importance overall relative to the proof-first version for one passage (M=2.53 for principle-first and 2.85 for proof-first), but the opposite was true for the second passage (M=2.85 for principle-first and 2.48 for proof-first).

Insert Table 4 and Figure 2 about here

Principle and proof: Rating revisions. The number of revisions made in the importance ratings were analyzed with a logistic regression including the variables of type, passage, and structure. The analysis indicated that the best-fitting model included the main effects of type and structure, and the type x structure interaction (χ^2=1.28, df=4, p<.86). The regression coefficients for this model are shown in Table 5. The type x structure interaction is summarized by Figure 2b, which shows the mean proportion of rating changes for principle and proof content presented in a principle-first and a proof-first structure. The mean proportion of ratings changed in the principle segment was .32 with the principle-first structure and .28 with the proof-first structure. For the proof segment, these means were .24 with the principle-first structure and .39 with the proof-first texts. A greater proportion of sentence ratings were changed when the text segment was presented first in the passage. This suggests that there is greater uncertainty as to the relative importance of information when it occurs early in the passage. Additionally, the main effect of structure shows that more changes were made overall when the proof occurred first. The means for the main effect of structure were .37 for the proof-first texts and .26 for the principle-first passages. This indicates that readers are better able to predict the relative importance of sentences as they are reading with the principle-first structure.

Insert Table 5 about here

Different information types: Mean importance ratings. A multiway frequency analysis was also run on the sentences classified according to the 4-type scheme used in Experiment 1: core principle sentences, core proof sentences, non-core proof sentences, and introductory sentences. As with the principle and proof data, the analysis indicated no difference between the original and revised data sets; therefore, only the data for the original ratings will be reported. The best-fitting model included the main effect of type and the structure x passage interaction (χ^2=38.05, df=45, p<.76). Sentence type did not interact with structure in this data set (i.e., the inclusion of the type x structure interaction did not significantly improve the fit of the model). The coefficients for this model are shown in Table 6.   The mean importance ratings for the four sentence types are: core principle: 1.58; core proof 1.88; non-core proof: 2.28; introductory: 3.45.   Readers considered the core principle sentence to be more important than the core proof sentence, even though they stated essentially the same information in different contexts, suggesting that readers consider the principle to be the more important category of information (i.e., the principle is judged more important when presented as a statement of the principle rather than the outcome of the proof). This importance judgement could partially account for why readers use the principle-first structure in summarizing both principle-first and proof-first passages. If readers think that important information should occur at the beginning of a text, and consider the core principle to be the most important information, they would mention the principle first when summarizing.

> Insert Table 6 about here

The structure x passage interaction indicates that readers rated the principle-first version of one passage higher in importance overall relative to its proof-first version; the opposite was true for the second passage. The mean importance ratings for the first passage were 2.52 for the principle-first version and 2.85 for the proof-first version. These means for the second passage were 2.85 for the principle-first version and 2.48 for the proof-first version.

Different information types: Rating revisions. The number of ratings revised for the four sentence types was analyzed with a logistic regression. There were significant effects due to type and structure, with no interactions (χ^2=16.92, df=11, p<.11). The coefficients for this analysis are given in Table 7. The proportion of ratings changed are: core principle: .12; core proof: .20; non-core proof: .36; introductory: .33. Readers changed the fewest ratings for the core principle sentences, indicating that they were able to accurately predict the importance of this content as they were reading. They thought that this was the most important content as they were reading, and did not change this assessment after finishing the text. The next fewest changes were in ratings of the core proof sentences, followed by the non-core proof sentences and introductory sentences, both of

which had relatively high revision rates (about a third of the ratings were changed). As in the previous analysis, the main effect of structure shows that more ratings were changed overall in the proof-first than principle-first passages.

```
Insert Table 7 about here
```

Experiment 3

The first two experiments showed that principle-first and proof-first structures produce differences both in attention and in perceived importance of the principle and proof content, with the information being judged as more important and processed more thoroughly when it was presented first. These studies also found that the principle-first structure was easier for readers to process (in terms of being able to accurately predict importance during reading), and that readers preferred to use this structure when summarizing. The purpose of the third study was to determine whether these processing differences and structural preferences are reflected in readers' memorial representation of the texts. In this study, subjects read passages having the two contrasting organizations and then completed a free-recall both immediately and after a one-week delay. The recalls were examined to determine whether memory for the principle and proof information varied according to the manner in which they were presented. Additionally, the order in which the principle and proof were recalled was analyzed to see if readers tended to recall the information in the preferred principle-first format, or in an organization corresponding to that of the passage they read.

Method

Materials. The materials were the same two passages used in Experiment 2. A minor change was made in one of the passages to tighten up the logic of the proof and make it more cohesive.

Subjects. The subjects were 26 undergraduates who had completed not more than one semester of college-level physics.

Procedure. The subjects were told that the purpose of the experiment was to determine what types of information are easiest to understand in scientific texts. They were told that they would be reading four short passages and taking a test on the contents of each immediately after reading each passage, and that the type of test they received would vary with the passage. Subjects read two practice passages and two target passages, with one practice passage preceding each target passage. For the target passages, subjects were told to write down everything they could recall in as close to the original form as possible. The tests for the two practice passages consisted of multiple-choice, matching, and short-answer questions. Different types of tests were used for the practice passages so subjects would not expect free-recall tests and therefore try to memorize the passages.

Subjects returned one week later expecting to read four new passages, but were told to write down everything they could recall from the two target passages that they had read the previous week.

Results

The free recalls were scored in the following ways. First, in order to determine whether subjects reorganized the passages, the recalls were scored according to whether or not the subject had recalled the gist of the principle and proof, and the order in which those two were recalled. Each recall was first scored for full or partial recall of the principle and proof. If both were judged to be recalled (either partially or in full), then the recall was scored for order in which these two were presented. The recalls were scored by two scorers, and there was 97% agreement on whether the principle and proof had been included, and the order of recall.

Second, in order to look at the amount recalled from the principle and proof, each recall was scored for the number of sentences and propositions recalled correctly and incorrectly. A correctly recalled sentence was defined as one in which all of the main propositions were recalled accurately. Main propositions were defined as propositions expressing the main idea of the sentence. Partial credit was given for recall of sentences having more than one main idea. To receive partial credit, the subject had to recall all of the main propositions for at least one idea. A correctly recalled proposition was either recalled verbatim or recalled with substitutions that represented the gist of the proposition elements. The free recalls were scored by two scorers to assess interscorer reliability. There was 94% agreement between the two scorers on the proposition scoring.

Evidence of structural reorganization. The results of the earlier summary data (Experiment 1) indicated that novices find the principle-first structure more natural in that they prefer to use that structure in organizing passage summaries. However, there was no evidence of structural preference in the passage recalls. If subjects recalled information from both the proof and principle, they tended to recall it in the same order in which it was read. This was true for both immediate and delayed recall. In the immediate recalls, a total of 25 (out of 26) subjects recalled some information from both the principle and proof with the principle-first structure; only two reversed the information order. For the proof-first structure, 22 (out of 26) subjects recalled some content from both text segments and three of these restructured the text. In the delayed recalls, for the principle-first structure, 15 subjects recalled some information from both text segments and three of these changed the presentation order. For the proof-first organization, 11 subjects recalled both segments and one changed the order. Chi-square tests on the number of recalls in which the order was changed and unchanged with the two structures were not significant for either the immediate or delayed recall sets.

The free recall results indicate that when readers are asked to recall the passages in their original structure, they are able to do this regardless of which of the two structures they have read. These results indicate that readers have internal passage representations containing information about the

original text structure. Thus it is not clear to what extent they reorganize proof-first material internally. It could be that readers store the content in its original form and reorganize when writing a summary. Alternatively they may reorganize the content for storage but in some way tag the internal representation with information about its original structure.

Recall of principle and proof over time. Although it is not clear how structure influences the organization of the text representation, structure does appear to effect the strength of the representation, as indicated by whether or not principle and proof content dropped out of recall over time. The recall data were examined to determine whether those subjects who were initially able to recall the gist of the principle and proof were still able to recall this after the one-week delay, and the extent to which the ability to this varied with the passage structure. The degree to which principle and proof recall decreased over time was examined by looking at the proportion of subjects initially recalling the principle and proof who did not recall it one week later (i.e., the number of subjects recalling the principle initially minus the number recalling it at delay, divided by the number recalling it initially). This ratio represents the percentage decrease in recall after a delay. This was calculated for the principle and proof content for each of the two organizational structures. For the principle segment, there was a 33% (8/24) decrease in gist recall with the principle-first structure, and a 58% (14/24) decrease with the proof-first structure. The difference between these proportions is significant (z=1.74, p<.05, one-tailed test). Thus there was a greater decrease in recall of the principle with the proof-first structure. For the proof segment, there was a 64% (9/14) decrease in gist recall with the principle-first structure, and a 65% (11/17) decrease with the proof-first structure. These two proportions are not significantly different. The two structures thus produced differences in the strength of the representation of the principle segment of the text; the gist of the principle segment was more likely to be remembered over time with the principle-first structure.

Recall of the statements of the principle (the core sentences). In each passage the information that would generally be considered to be the most important to recall is the statement of the principle itself. The principle was presented in the two core sentences in each passage (the core principle and the core proof sentences). In order to determine whether structure influenced recall of this particularly important content, the number of subjects recalling these two sentences was examined for both immediate and delayed recall for each of the two structures. These data were analyzed with a multiway frequency analysis performed on the cell frequencies. The best-fitting model included the main effect of session (immediate vs delayed recall) and the type x structure interaction (χ^2=5.51, df=6, p<.48). The coefficients for this model are presented in Table 8. The type x structure interaction is seen in Figure 3 showing the number of people recalling the core principle and core proof sentences with the two structures. Figure 3 indicates that recall of the principle was enhanced by the principle-first structure, because this structure (relative to the proof-first structure) greatly

increased the probability of recalling the core principle sentence. With the proof-first structure, recall of the core principle and core proof sentences were comparable to each other, and to core proof statement with the principle-first structure. These effects did not interact with session, indicating a consistent pattern in the immediate and delayed recalls. This again shows that the principle-first structure enhanced recall of the principle information.

The main effect of session reflects the fact that more people recalled the core sentences immediately than after a one-week delay. The coefficients for the main effect of type show that the principle was recalled by more people than was the proof. The small coefficient for the main effect of structure indicate that structure had little effect on the total number of subjects recalling one or both of the core sentences.

```
┌─────────────────────────────────────────────────────────┐
│  ┌────────────────────────────────────────────────────┐ │
│  │        Insert Table 8 & Figure 3 about here        │ │
│  └────────────────────────────────────────────────────┘ │
└─────────────────────────────────────────────────────────┘
```

Amount recalled from principle and proof. The free recalls were scored for number of propositions and sentences recalled correctly and incorrectly. The number of recall errors was very low. Subjects recalled incorrectly an average of one proposition and less than one sentence per passage. Because the number of errors was so low, the errors themselves were not analyzed. Instead, analysis of variance tests (ANOVAs) were performed only on total recall and accurate recall (total recall with errors excluded). Separate analyses were performed on the number of sentences recalled and the number of propositions recalled. All significance tests on the means were done using the Newman-Keuls test with $p < .05$.

The sentence recall data were analyzed with an ANOVA on the proportion of sentences recalled from the principle and proof segments by each subject. The variables included in the analysis were passage structure (proof-first/principle-first), sentence type (principle/proof), and session (immediate recall/delayed recall). Separate analyses were performed on the total proportion of sentences recalled, and the proportion of sentences recalled accurately. In both cases, the analyses indicate a main effect of session and a type x structure interaction. Because the results of the two analyses were identical, only the accurate recall data will be reported.

The type x structure interaction ($F(1,25)=6.82$, $MS_e =.084$, $p<.015$) reflected the fact that subjects recalled more from the principle segment when it was presented first in the passage. This interaction is shown separately for immediate and delayed recall in Figure 4a. Significantly less principle information was recalled with the proof-first than principle-first structure, but there was no difference in the amount of proof content recalled with the two structures. Additionally, with the principle-first structure, there was no difference in the proportion of sentences recalled from the principle ($M=.41$) and the proof ($M=.34$). However, with the proof-first structure, subjects recalled

significantly less of the principle content (*M*=.27) than the proof content (*M*=.41). This type x structure interaction did not interact with session, indicating that the effect persists over time, as shown in Figure 4a. The main effect of session (*F*(1,25)=80.29, *MS*$_e$ =.029, *p*<.001) was due to the fact that subjects recalled more immediately (*M*=.46) than after a 1-week delay (*M*=.25).

Insert Figure 4 about here

ANOVAs were also performed on the proportion of propositions recalled from principle and proof sentences. Separate analyses were performed on the total proportion of propositions recalled and the proportion of propositions recalled accurately. Both analyses indicated significant main effects of session and structure, and structure x session and type x structure interactions. Additionally, the accurate recall data set contained a significant main effect of sentence type. Because the results of the two analyses were similar, only the accurate recall data will be presented.

The type x structure interaction in proposition recall (*F*(1,25)=9.77, *MS*$_e$ =.025, *p*<.004) reflected the same recall pattern as the corresponding interaction in sentence recall. The mean proportion of propositions recalled from the principle was .22 with the principle-first structure and .12 with the proof-first structure. The proportion of propositions recalled from the proof was .19 with the principle-first structure and .23 with the proof-first structure. As in sentence recall, subjects recalled significantly more of the principle propositions when the principle was presented first. Additionally, a significantly smaller proportion of principle than proof content was recalled in the proof-first passages. This interaction did not vary with session, as shown in Figure 4b.

The main effect of structure (*F*(1,25)=4.58, *MS*$_e$ =.015, *p*<.04) was due to the fact that subjects recalled more overall from the principle and proof paragraphs with the principle-first than proof-first passages. However, the structure x session interaction (*F*(1,25)=7.17, *MS*$_e$ =.005, *p*<.01) indicates that this was only true in the immediate recalls. In immediate recall, subjects recalled significantly more from the principle-first (*M*=.28) than proof-first (*M*=.22) texts, but overall recall with these two organizations did not differ after a one-week delay (*M*=.14 for principle-first; *M*=.13 for proof-first).

The main effect of session (*F*(1,25)=79.04, *MS*$_e$ =.009, *p*<.001) was due to the fact that subjects recalled more immediately (*M*=.25) than after a week delay (*M*=.13). The main effect of type (*F*(1,25)=4.10, *MS*$_e$ =.019, *p*<.054) was due to the fact that subjects recalled a greater proportion of propositions from the proof paragraph (*M*=.21) than principle paragraph (*M*=.17).

Summary. The free recall data indicates that organization influenced the ability to recall the principle. Readers were more likely to recall the gist of the principle segment after a delay when they had read the principle-first texts. Additionally, more readers recalled the core principle sentence with the principle-first than proof-first texts, with no significant difference in recall of the core proof

sentence. Finally, the amount of information recalled from the principle portion of the text was greater with the principle-first structure. This latter effect in amount recalled was reflected in the data in the following two ways.

First, readers recalled a greater proportion of the information in the principle segment with the principle-first than proof-first structure. There was not a significant difference in the amount recalled from the proof with the two structures. Thus relative to the proof-first structure, the principle-first structure enhanced recall of the principle and did not influence the level of recall of the proof.

Second, a similar pattern is seen in the relative recall of the two types of content (principle vs. proof) within each structure. There was no difference in the proportion recalled from the principle and proof segments with the principle-first text, but subjects recalled less principle than proof content with the proof-first texts. Thus the proof-first structure decreased amount recalled from the principle relative to the proof; recall of the two was equal with the principle-first structure.

These effects are apparent both immediately and after a delay. The only interaction in the recall data between time of recall and text organization occurred in the number of propositions recalled. Subjects initially recalled more information overall with the principle-first organization, but this difference disappeared after a delay. The fact that this interaction was not apparent in the sentence data suggests that the locus of the effect is in recall of the less important content (i.e., information not considered to be part of the main ideas of the text sentences). This is information that was not considered in scoring for recall at the sentence level.

Discussion

This research examined the effects on text processing and recall of a specialized text structure commonly used in science and mathematics texts to present principles and their associated proofs. This was done by comparing the processing of this structure to that of an alternative organization more typical of those found in expository writing. The results indicate that these two structures produce differences in how principle and proof content is processed, organized, and recalled.

Processing Differences

Readers spend more time on the principle and proof information when this content is presented first, suggesting that this information is processed more thoroughly when it occurs at the beginning of a text. This finding is consistent with serial position effects that are typically found in reading times (Cirilo & Foss, 1980; Graesser, Hoffman, & Clark, 1980; Haberlandt & Graesser, 1985; Olson, Mack, & Duffy, 1981). This effect has been attributed to the fact that the beginning of a text usually contains more new information (e.g., new objects, concepts, terms) than the middle or end, and encoding new concepts into the text representation requires more time (Graesser, Hoffman, & Clark, 1980). Similarly, it has also been argued that this effect is due in part to the fact that it is easier to integrate successive sentences into the reader's evolving model of the text topic as the model is

progressively refined (Haberlandt & Graesser, 1985). The serial position effect in reading times is almost certainly not a single effect, but the result of several different kinds of processing that occur when a reader is progressing through a text. The results of the current research suggest that the extra processing time spent on the content at the beginning of a text is due in part to (a) the increase in the perceived importance of the content when it occurs first, and (b) the difficulty in determining the precise relative importance of beginning content.

Perceived importance. Readers judged information to be more important when it was presented first. Thus the proof content was considered to be more important when in the proof-first structure, and the principle content judged more important in the principle-first structure. This finding is consistent with research which has shown that readers assume that the information presented first in a passage is particularly important (Kieras, 1979, 1985). Readers use initial placement as an indicator of importance, and in particular expect thematic information to appear in this position (Kieras, 1980a). This effect of information order on perceived importance could be partly responsible for the corresponding order effect found in the reading times in this research. Readers may have spent more time processing the principle and proof when they occurred first because readers thought the initially presented content would be particularly important.

Ability to predict importance. This research also found that readers were less able to assess the relative importance of information when it was presented first. In the importance ratings task, more ratings were changed (in both principle and proof) when they occurred first, indicating that readers were less able to predict the importance of the beginning content. Therefore the additional time spent on principle and proof content when they occur first may be due in part to readers' uncertainty about the relative importance of that content.

This is consistent with explanations of serial position effects in reading time that conceptualize comprehension as a process of progressive model refinement (Haberlandt & Graesser, 1985). This hypothesis assumes that readers begin with a general text topic and build a model of the text content by relating subsequent information to that topic and thus gradually elaborating the model (Collins, Brown, & Larkin, 1980). The model building processes occurring at the beginning of the text are relatively demanding (the reader must activate relevant knowledge structures and establish new structures), but integrating new information into the representation becomes easier as the representation grows and there is more overlap between the new text content and the existing representation (Haberlandt & Graesser, 1985; Olson, Mack, & Duffy, 1981; Townsend, 1983). Thus reading times are faster for information occurring later in the text because sentence integration is easier.

The results of the current research indicate that although readers expected the initial content in a passage to be important, they had difficulty predicting its precise importance, presumably because

they had only a rudimentary model for the text at that point (i.e., possibly a general idea as to the topic based on the title and introductory material). In contrast, when the content appeared later in the text, the text representation was almost complete and the status of the content within that representation was clear. Thus readers were better able to predict the importance of content when it appeared at the end.

Ease of comprehension. In addition to the processing differences found for proof and principle content with the two structures, it was expected that novices would find the principle-first structure easier to process overall. This is because it is a more familiar organization and provides the reader with an initial conceptual framework for processing the text. This expectation would be reflected in main effects of organization on reading times and on readers' ability to predict importance. The two structures did not produce differences in overall reading times, but did produce differences in predicting importance. Readers had particular difficulty predicting the importance of information with the proof-first structure. There were more on-line sentence ratings changed after reading with the proof-first than principle-first passages. This suggests that subjects are able to use the principle (when it is presented first) as an conceptual framework and determine what is important in the proof by assessing its relationship to the principle. It is apparently less clear what is important in the proof when the principle (i.e., what is ultimately being proven) is not known.

This finding is consistent with text comprehension theories emphasizing the role of superordinate information in guiding text processing. In the van Dijk and Kintsch (1983; Kintsch & van Dijk, 1978) processing model, readers build text representations at several levels by generating or selecting from the text initial information units which function as superordinates to which the subsequent information is linked. At a local level, readers attempt to relate incoming propositions to a superordinate proposition in a way that maintains the coherence of the text, and in this way form a propositional textbase representing in detail all of the content of the passage. Simultaneously, at a global level, readers try to generate a macrostructure representing the main points of the passage. This is done by inferring from the propositional textbase a set of macropropositions capturing the passage gist. The ease of this gobal processing depends in part on cues indicating what content expresses thematic information or provides information from which thematic information can be generated. One technique for facilitating this macrolevel processing is to include early in the text thematic statements presenting main idea information. These statements can then be used as a superordinate framework guiding processing of subsequent information (see van Dijk & Kintsch, 1983, pp 201-203 for discussion).

In the proof-first structure, there are few thematic cues facilitating macrolevel processing. The overall structure is unfamiliar (so the novice reader does not expect certain types of content to occur at certain places in the text) and there is no substantive thematic statement. Thus there is little to

guide the reader in constructing the text macrostructure, other than the title and general topic information in the introduction. Accordingly, readers are less able to assess the importance of information as they are reading. With the principle-first structure, the text provides the reader at the beginning with a thematic statement (i.e., the principle). This statement apparently facilitates readers' ability to assess the importance of the text content.

The relative inability of subjects to accurately predict importance with the proof-first structure suggests that readers experience more confusion as they are reading, and thus find this structure more difficult to process. However, there is not direct emprical evidence for this hypothesis, as there was no difference in overall reading time for the two passages. In general, research with passages having structural variations similar to those in the current study has reported mixed results on whether these structural differences influence reading times (Kieras, 1980a). There are several reasons why an increase in uncertainty might not be reflected in reading times in this study. It is possible that the level of uncertainty about importance produced by the proof-first organization was simply not large enough to be reflected in reading times at the sentence level. Additionally, the effect may have been localized to certain sentences in the text, and was therefore washed out by the inclusion of all sentences in the reading time analyses.

It is also possible that the readers were not less certain as to the correctness of their ratings while they were reading the proof-first texts, but changed their minds about the relative importance of the content after reading the principle at the end, and thus did not experience more comprehension difficulty as they were reading (i.e., they were confident of the incorrect ratings until reaching the principle segment of the passage). However, this hypothesis would also predict an increase in processing time for the proof-first passage, in that readers would at some point realize that they had assumed an incorrect interpretation, and reorganize their internal representation of the text accordingly. This reorganization would be expected to add to the overall processing time for the proof-first structure. Thus the lack of a difference in overall reading time does not support the hypothesis that readers' were confident of their incorrect ratings while reading the proof-first texts.

Information Organization

The summary data showed that readers' clearly preferred the principle-first structure, in that almost all subjects used this organization in summarizing. This result suggests that readers thought this organization was most appropriate for the passage content. This finding is similar to that of Kintsch, Mandel, and Kozminsky (1977), who found that subjects summarizing stories whose paragraphs had been randomly reordered always restructured the stories into a canonical ordering, so their summaries matched those of subjects reading well-formed stories. The fact that subjects always placed the principle first in their summaries suggests that novices thought the statement of the principle was the most important content in the passage. This was also indicated by the ratings data

(i.e., the core principle statement was rated highest in importance independent of the passage structure).

There are two reasons why the principle would be perceived by readers as the most important text content, even when it occurred last. First, as noted earlier the principle statement provided a superordinate conceptual framework for organizing the rest of the passage content. It is the only statement in the passages that is related to all of the other content. Thus it is reasonable for subjects to give this content primary status in text representation. Second, the principle is the only information in the passage that is repeated. It has been shown that readers assume that repeated information is particularly important (Perfetti & Goldman, 1974).

Although subjects restructured the proof-first passages in their summaries, they tended to maintain the original passage structure in their recalls, even after a one-week delay. Thus subjects who recalled both principle and proof information were capable of recalling the original form of the passages when requested to do so. This indicates that organization information is present in novices' internal representation of the text. Therefore, it is not clear whether subjects reorganized the passages internally. They may store the textbase in its original form (and reorganize it for purposes of summarizing), or they may reorganize the textbase but mark the representation as to its initial structure so this information is available when needed.

Recall of Principle and Proof Information

The processing differences resulting from the two structures influenced memory for the principle, but did not affect recall of the proof. In the case of the proof, readers were just as likely to recall the gist of the proof after a delay with the proof-first structure as with the principle-first structure. Additionally, the amount recalled from the proof did not vary with text structure. Finally, the number of readers recalling the particularly important core proof statement (i.e., the statement of the principle as the outcome of the proof) did not differ for the two structures. Thus presenting the principle before or after its corresponding proof does not substantially alter the internal representation formed for the proof, even though presenting the principle first facilitates the on-line processing of the proof by increasing its predictability.

The absence of a strong effect of structure on proof recall may be due to the fact that the proof is completely understandable as a stand-alone unit, and can be processed independently of the principle segment of the text. As noted earlier, the proof-first passages found in scientific and mathematics texts are comprehensible to readers, but potentially difficult to process. Thus the proof can be understood without the principle portion of the text. Therefore this content is equally memorable with both structures, even though its processing may be facilitated by the principle-first structure.

On the other hand, the text structure did affect recall of the principle portion of the text. Readers were more likely to recall long-term the principle segment with the principle-first structure. They also recalled more from the principle segment with this structure. Finally, the number of readers recalling the core principle sentence (i.e., the statement of the principle) was greater with the principle-first structure. Thus the principle-first structure facilitated recall of the principle portion of the text, both immediately and after a delay.

These recall differences for the principle segment of the text could result from two processing differences between structures. First, the facilitory effect of the principle-first structure on recall of the principle segment of the text could be due to the fact that this structure permits readers to use the principle as a conceptual framework for processing the text as they read. The summary data indicates that readers with both structures ultimately give the principle primary status in their macrostructure for the text. However, the principle-first structure allows readers to form this macrostructure initially, while the proof-first structure requires that readers generate this structure later in the reading of the passage. According to the Kintsch and van Dijk processing model, the information units that readers use as superordinates to guide their text processing are repeatedly re-processed as readers try to relate subsequent text information to the superordinate concepts. Because superordinate units are processed more often, they are better recalled (van Dijk & Kintsch, 1983, pp 44-45). Thus the principle may be better recalled with the principle-first structure because it is processed more frequently (i.e., repeatedly held in working memory) as readers use the principle to interpret the proof.

Second, the enhanced recall of the principle segment with the principle-first structure may reflect differences in the emphasis given to that segment due to the spatial contiguity of the core principle and core proof statements in the two structures. In the proof-first structure, the principle segment follows almost immediately the core proof statement of the principle, and therefore provides information redundant with that just read (i.e., the core principle statement is a restatement of the core proof sentence). Readers may therefore simply view the principle segment as review material, and not process it as thoroughly, such that less content is recalled from that segment and it is forgotten over time. Thus even though the core principle statement is viewed as important with this structure (in sentence ratings and summaries), it is not well-processed or strongly linked to the rest of the passage in memory because it is redundant with preceding content in the proof. The principle is therefore recalled as part of the proof, but not as a separate portion of the text in the principle segment.

On the other hand, in the principle-first structure, the principle is first presented as an independent unit, and then again as part of the proof at the end. Thus it may be processed more thoroughly when presented as part of the principle segment because this is its first presentation.

Additionally, it is reviewed again at the end of the text in the context of the proof, providing for spaced review of the content. Therefore readers recall the principle both as an independent unit, and as part of the proof. The principle-first structure thus increases the probability of readers remembering the principle because they are more likely to recall both presentations of the principle in the text, recalling it as an independent fact and as the outcome of the proof.

The principle-first structure also resulted in subjects recalling more propositions overall immediately after reading the passage. However, this effect not found either in sentence recall or at delay. This indicates that the effect was localized to less important text content (information not considered to be central in the passage sentences). This suggests that the principle-first structure may have provided readers with a tighter organizational framework which enhanced recall of content incidental to the main points of the principle and proof. However, because this content was not critical to understanding the passage as a whole, it probably was not strongly linked to the rest of the passage in memory, and was rapidly forgotten, such that overall recall was equal for both structures after a one-week delay.

Summary

This research contrasted the processing and recall of two expository text structures which differ in their schematic familiarity and in the availability of superordinate concepts for organizing on-line text processing. The results indicated that novice readers find the specialized proof-first structure less appropriate and more difficult to process. This organization also decreases the probability of readers recalling the passage principle. These results suggest that the proof-first structure increases the difficulty of developing an appropriate text macrostructure, and results in a more fragmented representation, such that the principle segment is poorly recalled.

The results of this research suggest that the structure used by a writer in presenting this type of information should depend on the instructional goals of the text. If the main goal is to have the learner understand the principle, and the proof is provided as additional elaborative information designed to strengthen understanding of the principle, then the principle-first structure seems most appropriate. This structure adheres to the writing recommendations suggested by text processing principles, and, as indicated by the current research, results in the desired learning outcomes. Additionally, this structure corresponds to novices' view of how this type of content should be organized, most likely because it is consistent with writing conventions regarding the placement of topic information. In this sense, it corresponds to an expository text schema that novices have for this type of content.

Understanding the principle is the most common learning goal associated with this type of scientific content, and is probably the goal that is most often associated even with proof-first passages in existing science texts. However, other goals are possible for which the proof-first

structure might be more appropriate. One would be if the main topic of interest were the methodology underlying the proof; another would be if the logic behind the proof were of most concern. In these cases the nature of the proof would be of greater interest than the specific principle. Thus a structure analogous to an argument leading up to a conclusion would be most appropriate in that it would focus attention on this content. However, the current research indicates that in these cases, providing an informative introduction to the proof containing appropriate orienting information would help readers process texts having this type of structure.

References

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grune and Stratton.

Bransford, J. D., & Johnson, M.K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383-438). New York: Academic Press.

Cirilo, R., & Foss, D. (1980). Text structure and reading time for sentences. *Journal of Verbal Learning and Verbal Behavior,* **19,** 96-109.

Collins, A. M., Brown, J.S., & Larkin, K.M. (1980). Inference in text understanding. In R.J. Spiro, B. C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 385-407). Hillsdale, NJ: Lawrence Erlbaum.

Danet, B. (1980). Language in the legal process. *Law & Society Review,* **14,** 445-564.

Fienberg, S. E. (1980). *The analysis of cross-classified categorical data.* Cambridge, MA: MIT Press.

Graesser, A. C., Hoffman, N.L., & Clark, L.F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior,* **19,** 131-151.

Haberlandt, K. F., & Graesser, A.C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General,* **114,** 357-374.

Kieras, D. E. (1979). *Initial mention as a cue to the main idea and main item of a technical passage* (Tech. Rep. 3). University of Arizona.

Kieras, D. E. (1980a). Initial mention as a signal to thematic content in technical passages. *Memory & Cognition,* **8,** 345-353.

Kieras, D. E. (1980b). *Abstracting main ideas from technical prose: A preliminary study of six passages* (Tech. Rep. 5). University of Arizona.

Kieras, D. E. (1985). Thematic processes in the comprehension of technical prose. In B. K. Britton and J. B. Black (Eds.), *Understanding expository text* (pp. 89-107). Hillsdale, NJ: Lawrence Erlbaum.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review,* **85,** 363-394.

Kintsch, W., & Yarbrough, J.C. (1982). The role of rhetorical structure in text comprehension. *Journal of Educational Psychology,* **74,** 828-834.

Kintsch, W., Mandel, T.S., & Kozminsky, E. (1977). Summarizing scrambled stories. *Memory & Cognition,* **5,** 547-552.

Mandler, J. M. (1978). A code in the node: The use of a strong schema in retrieval. *Discourse Processes,* **1,** 14-35.

Meyer, B. J. F. (1983). Text dimensions and cognitive processing. In H. Mandl, N. Stein, & T. Trabasso (Eds.), *Learning and understanding from text* (pp. 3-47). Hillsdale, NJ: Lawrence Erlbaum.

Olson, G. M., Mack, R.L., & Duffy, S.A. (1981). Cognitive aspects of genre. *Poetics*, **10**, 283-315.

Perfetti, C. A., & Goldman, S.R. (1974). Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior*, **13**, 70-79.

Townsend, D. J. (1983). Thematic processing in sentences and texts. *Cognition*, **13**, 223-261.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Vesonder, G. T. (1979). *The role of knowledge in processing experimental reports.* Unpublished doctoral dissertation, University of Pittsburgh.

Table 1

Sample Passage Showing a Proof First and Principle First Organization

Archimedes' Principle - The Buoyant Force

Buoyancy is a familiar phenomenon. For example, a body seems to weigh less when it is immersed in water than when it is surrounded by air. This is due to the buoyant force exerted by the water on the immersed body. The water exerts pressure on the body's surface, causing an upward force on the bottom and a downward force on the top of the object. Because the pressure is greater at greater depths, the upward force will be greater and the resultant force will be an upward or "buoyant" force.

[Continuation of proof-first version]
Consider an arbitrary portion of a fluid at rest. This portion of fluid occupies a volume V with a surface S. (Diagram showing a labeled volume V with surface area S). This fluid volume V is held at rest by the forces exerted by the remainder of the fluid on the surface S. (Diagram showing arrows directed from the surrounding fluid towards the volume V.) Because the fluid portion is at rest, the upward buoyant force must exactly balance the downward gravitational force. Therefore the buoyant force must be directed vertically upward and be equal in magnitude to the weight of the fluid in V. Assume that we now replace the fluid inside V with an object of the same shape and size. The pressure at every point on surface S is exactly the same as before; the buoyant force exerted on the body by the surrounding fluid is unaltered. The buoyant force must therefore be equal to the weight of the displaced fluid. The weight of the displaced fluid is the gravitational constant g times the mass, ρV, where ρ is the density of the fluid. Thus, the buoyant force on the object is equal to the product

$$\rho V g.$$

This is Archimedes' principle for finding the magnitude of a buoyant force.

This very simple expression for determining the magnitude of a buoyant force was discovered by Archimedes over 2000 years ago. Archimedes' principle states that the buoyant force B exerted on an object immersed in a fluid is

$$B = \rho V g,$$

where V is the volume of a displaced homogeneous fluid of density ρ, and g is the gravitational constant. This result can be deduced from the laws of fluid statics, as shown above.

[Continuation of principle-first version]
A very simple expression for determining the magnitude of a buoyant force was discovered by Archimedes over 2000 years ago. Archimedes' principle states that the buoyant force B exerted on an object immersed in a fluid is

$$B = \rho V g,$$

where V is the volume of a displaced homogeneous fluid of density ρ, and g is the gravitational constant. This result can be deduced from the laws of fluid statics, as shown below.

Consider an arbitrary portion of a fluid at rest. This portion of fluid occupies a volume V with a surface S. (Diagram showing a labeled volume V with surface area S). This fluid volume V is held at rest by the forces exerted by the remainder of the fluid on the surface S. (Diagram showing arrows directed from the surrounding fluid towards the volume V.) Because the fluid portion is at rest, the upward buoyant force must exactly balance the downward gravitational force. Therefore the buoyant force must be directed vertically upward and be equal in magnitude to the weight of the fluid in V. Assume that we now replace the fluid inside V with an object of the same shape and size. The pressure at every point on surface S is exactly the same as before; the buoyant force exerted on the body by the surrounding fluid is unaltered. The buoyant force must therefore be equal to the weight of the displaced fluid. The weight of the displaced fluid is the gravitational constant g times the mass, ρV, where ρ is the density of the fluid. Thus, the buoyant force on the object is equal to the product

$$\rho V g.$$

This is Archimedes' principle for finding the magnitude of a buoyant force.

A body whose average density is less than that of a liquid's density can float partially submerged in the liquid. Such a body sinks until its weight is balanced by the buoyant force. According to Archimedes' principle, this buoyant force is equal to the weight of the displaced liquid. Hence the weight of the displaced liquid is equal to the weight of the floating body.

Table 2

Regression Model of the Log of the Sentence Reading Times (Milliseconds)[1]

Variable	Coeff.	St. Error	St. Reg. Coeff.	P-value
Intercept	3.350			
Session	.056	.015	.093	.001
Words	.022	.001	.643	.001
Passage				
Passage 1	.038	.010	.099	.001
Passage 2	.029	.011	.082	.006
Passage 3	-.067	.010	-.186	.001
Structure	-.026	.017	-.043	.119
Type	-.124	.023	-.178	.001
Type x Structure	.101	.032	.112	.002

[1]R^2 = .51; multiple R = .71

Table 3

Regression Model of the Log of the Sentence Reading Times (Milliseconds)[2]

Variable	Coeff.	St. Error	St. Reg. Coeff.	P-value
Intercept	3.430			
Session	.056	.014	.093	.001
Words	.017	.001	.488	.001
Passage				
Passage 1	.047	.010	.123	.001
Passage 2	.012	.011	.033	.253
Passage 3	-.059	.010	-.165	.001
Structure	.013	.017	.022	.436
Type				
Core Principle	-.053	.026	-.084	.038
Core Proof	.023	.024	.037	.334
Non-core Proof	.103	.015	.224	.001
Introductory	-.073	.021	-.148	.001
Type x Structure				
Core Principle	.111	.031	.124	.001
Core Proof	-.059	.034	-.066	.082
Non-core Proof	-.037	.021	-.061	.087
Introductory	-.015	.023	-.022	.525

---

[2] R^2 = .54; multiple R = .74

Table 4

Loglinear Model of the Immediate Sentence Importance Ratings

	Variable	Coeff.	St. Error	Coeff./St. Error
(a)	Structure: (Principle-1st estimates)			
	Rate 1:	.071	.101	.70
	Rate 2:	-.216	.126	-1.71
	Rate 3:	.116	.102	1.14
	Rate 4:	.002	.109	.02
	Rate 5:	.026	.136	.19
(b)	Type: (Principle estimates)			
	Rate 1:	.047	.100	.47
	Rate 2:	-.461	.123	-3.74
	Rate 3:	-.123	.101	-1.21
	Rate 4:	.113	.103	1.10
	Rate 5:	.424	.130	3.26
(c)	Type x Structure: (Estimates for proof in principle-1st structure)			
	Rate 1:	.099	.099	1.00
	Rate 2:	-.156	.124	-1.26
	Rate 3:	.273	.101	2.69
	Rate 4:	-.050	.104	-.48
	Rate 5:	-.167	.130	-1.28
(d)	Passage: (Passage 1 estimates)			
	Rate 1:	.053	.090	.59
	Rate 2:	-.085	.089	-.95
	Rate 3:	-.050	.085	-.59
	Rate 4:	.200	.097	2.05
	Rate 5:	-.118	.132	-.89
(e)	Structure x Passage: (Estimates for proof in passage 1)			
	Rate 1:	.036	.088	.41
	Rate 2:	.357	.089	4.01
	Rate 3:	.073	.084	.87
	Rate 4:	-.167	.097	-1.71
	Rate 5:	.300	.133	-2.26

Table 5

Logistic Regression Model of the Number of Revised Sentence Importance Ratings

Variable	Coeff.	St. Error	Coeff./St. Error
Type:	.035	.115	.30
Structure:	.135	.115	1.17
Type x Structure:	.215	.115	1.86

Table 6

Loglinear Model of the Immediate Sentence-Importance Ratings

	Variable		Coeff.	St. Error	Coeff./St. Error
(a)	Type:				
	Rate 1:	Core Principle	1.038	.206	5.05
		Core Proof	.634	.215	2.95
		Non-core Proof	.109	.158	.69
		Introductory	-1.781	.226	-7.89
	Rate 2:	Core Principle	-.271	.279	-.97
		Core Proof	-.009	.265	-.03
		Non-core Proof	.616	.162	3.79
		Introductory	-.336	.172	-1.95
	Rate 3:	Core Principle	-.473	.292	-1.62
		Core Proof	.039	.250	.16
		Non-core Proof	.109	.171	.64
		Introductory	.325	.159	2.04
	Rate 4:	Core Principle	-.422	.344	-1.23
		Core Proof	-.293	.322	-.91
		Non-core Proof	.026	.201	.13
		Introductory	.688	.183	3.76
	Rate 5:	Core Principle	.128	.405	.32
		Core Proof	-.372	.463	-.80
		Non-core Proof	-.059	.333	-2.58
		Introductory	1.103	.239	4.62
(b)	Structure: (Principle-1st estimates)				
	Rate 1:		.037	.088	.42
	Rate 2:		-.074	.087	-.85
	Rate 3:		-.016	.082	-.20
	Rate 4:		.024	.096	.25
	Rate 5:		.030	.126	.24
(c)	Passage: (Passage 1 estimates)				
	Rate 1:		.050	.087	.58
	Rate 2:		-.083	.087	-.95
	Rate 3:		-.049	.083	-.59
	Rate 4:		.191	.095	2.01
	Rate 5:		-.109	.126	-.87
(d)	Structure x Passage: (Estimates for principle-1st structure in passage 1)				
	Rate 1:		.031	.087	.36
	Rate 2:		.341	.087	3.92
	Rate 3:		.067	.083	.80
	Rate 4:		-.163	.095	-1.72
	Rate 5:		-.275	.126	-2.19

Table 7

Logistic Regression Model of the Number of Revised Sentence-Importance Ratings

Variable	Coeff.	St. Error	Coeff./St. Error
Type			
Core Principle:	-.800	.377	-2.12
Core Proof:	-.236	.325	-.72
Non-core Proof:	.585	.195	3.00
Introductory:	.450	.196	2.30
Structure:	.249	.101	2.48

Table 8

Loglinear Model of the Number of Subjects Recalling the Core Sentences

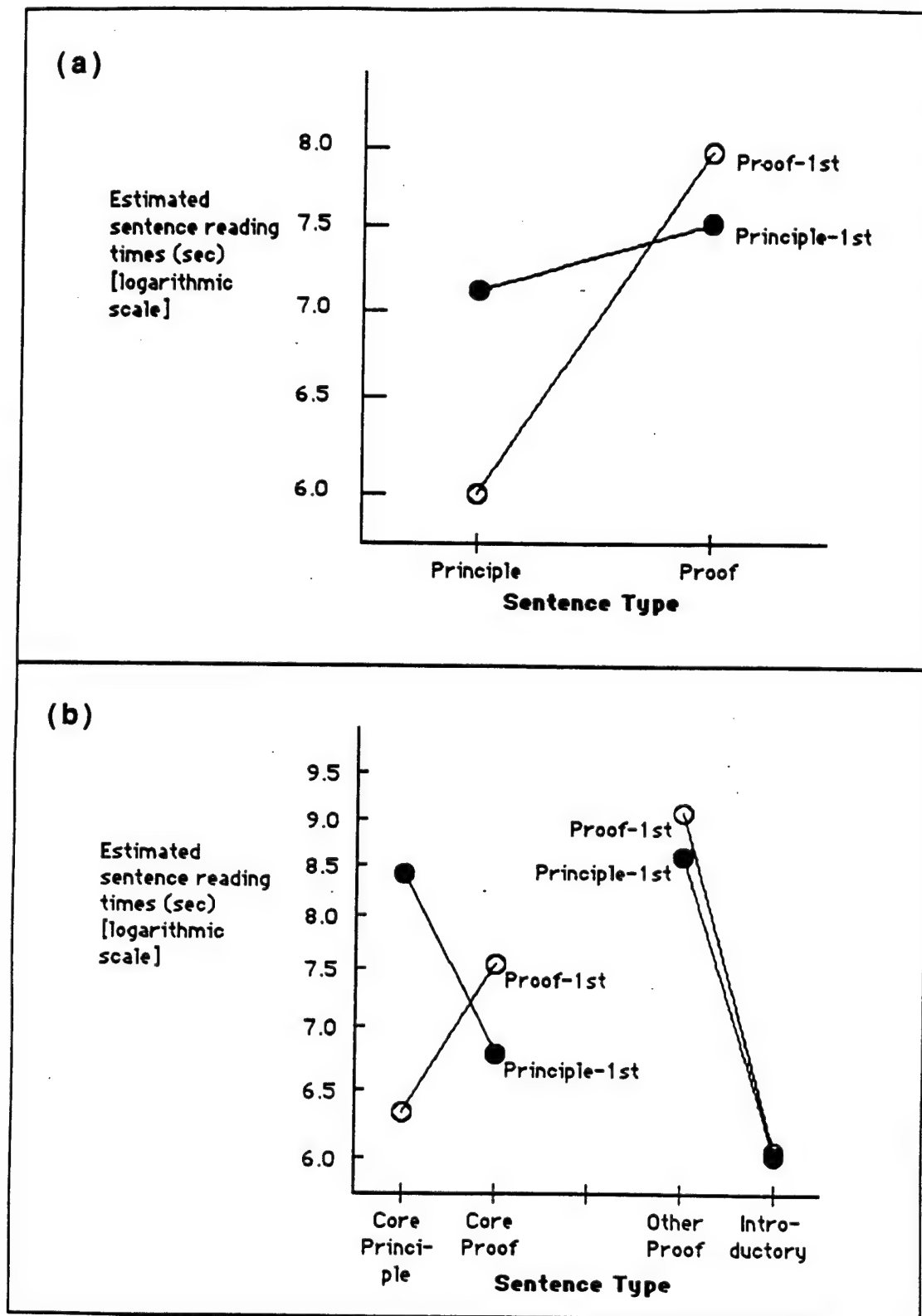|	Variable	Coeff.	St. Error	Coeff/St. Error
(a)	Session: (Immediate recall estimates)	.376	.079	4.76
(b)	Structure: (Principle-1st estimates)	.057	.077	.74
(c)	Type: (Principle estimates)	.188	.077	2.44
(d)	Type x Structure: (Estimates for principles in principle-1st structure)	.257	.077	3.33

Figure Captions

Figure 1. Estimated sentence reading times for type by structure interactions. (a) Principle and proof segments.   (b) Core principle, core proof, non-core proof, and introductory sentences.

Figure 2.   The type x structure interactions in the ratings data.   (a) Mean importance ratings. (b) Proportion of ratings changed.

Figure 3. Number of readers recalling the two core sentences.

Figure 4. The type x structure interactions in amount recalled. (a) Proportion of sentences recalled. (b) Proportion of propositions recalled.

Figure 1: Estimated sentence reading times for type by structure interactions.
(a) Principle and proof segments. (b) Core principle, core proof, non-core proof, and introductory sentences. ‑
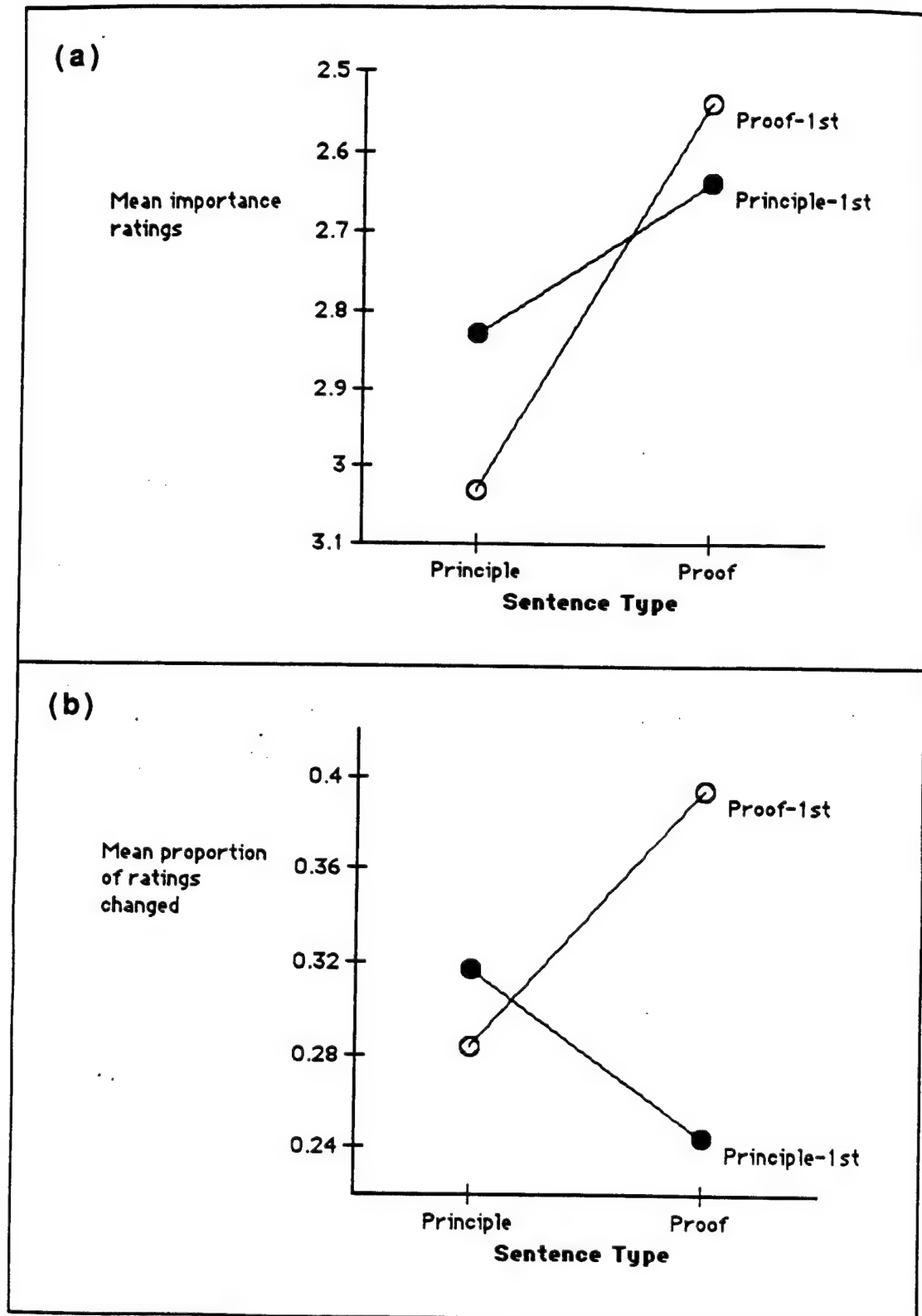
Figure 2: The type x structure interactions in the ratings data. (a) Mean importance ratings.
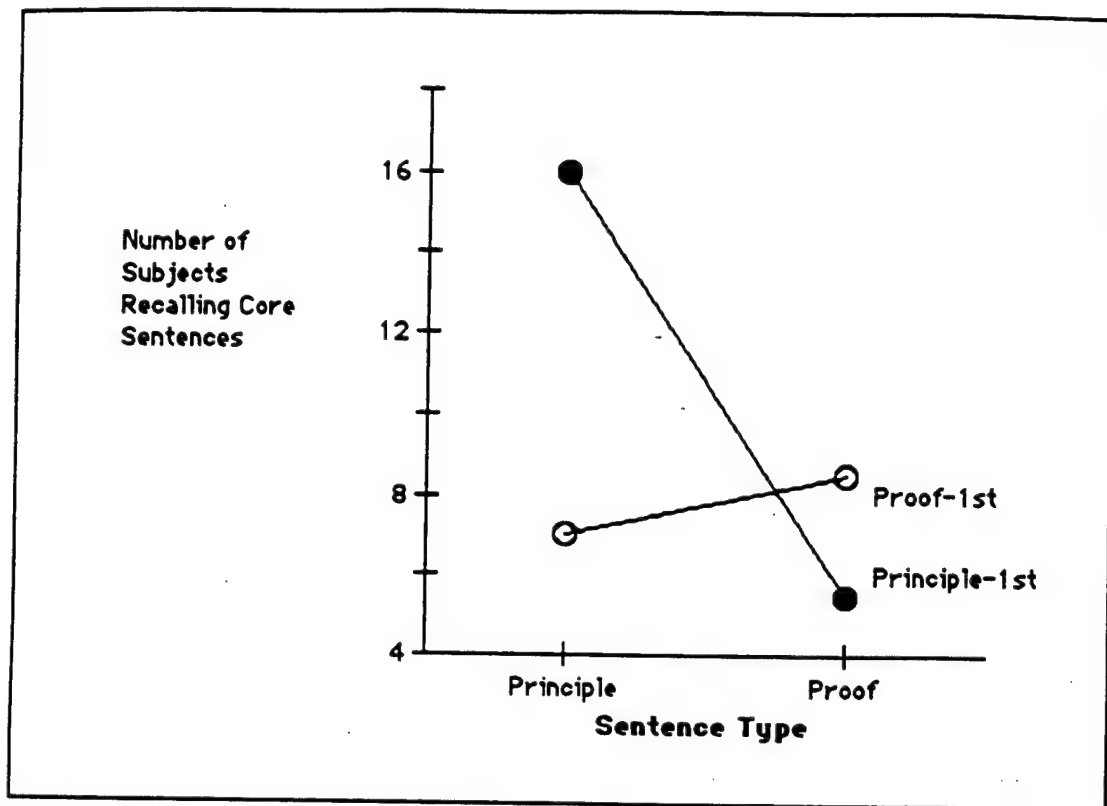(b) Proportion of ratings changed.

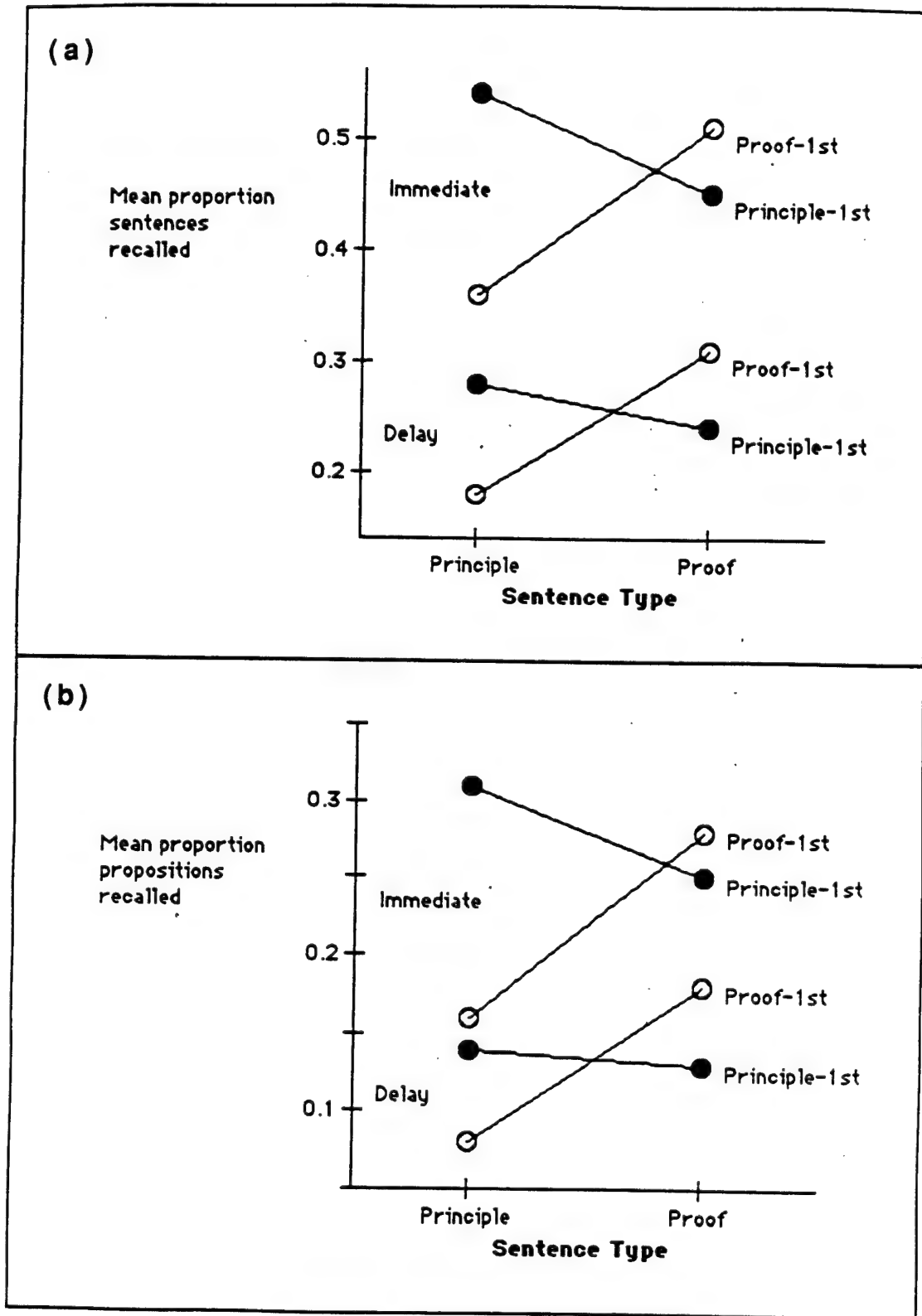Figure 3: Number of readers recalling the two core sentences.

Figure 4: The type x structure interactions in amount recalled. (a) Proportion of sentences recalled. (b) Proportion of propositions recalled.

Appendix D

(Total pages in length: 22)

Dee-Lucas, D. & Larkin, J.H. (1989b).

Equations in scientific proofs: Effects on comprehension.

(Tech. Rep. A.C.P. 20). Carnegie Mellon University

Equations in Scientific Proofs: Effects on Comprehension[*]

Diana Dee-Lucas & Jill H. Larkin
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

A.C.P. #20

August 1989

Running head: Equations in Proofs

Abstract

Scientific laws are typically presented with their corresponding proofs to justify their correctness and to illustrate useful reasoning techniques. The form of these proofs is commonly a series of logically-related equations. This research examined how this equation-based format influences proof comprehension. This was done by comparing readers' ability to solve problems after studying typical equation-based proofs and matching proofs with equations replaced by verbal equivalents (a verbal format). The equation-based format hindered proof comprehension by both reducing the comprehensibility of quantitative relations and decreasing attention to non-quantitative content. There was, however, no evidence that the deeper understanding of the proof provided by the verbal format influenced comprehension of the principle. The discussion addresses the role of proofs in scientific texts and how proofs should be presented.

Equations in Scientific Proofs: Effects on Comprehension

A major goal of instruction in science and mathematics is to teach fundamental laws and principles, and to provide an understanding of why these laws hold true. In support of this goal, students are often given proofs substantiating the relations summarized by a law that can be logically derived from other information. The proof format typically consists of a hypothetical situation (e.g., some objects in a general experimental setting) in which some variables are changed, and consequent changes in other variables are derived (e.g., how the temperature of a gas changes with an increase in external pressure). The relationships among these changes are often expressed as a series of equations. The initial equation expresses the relations among objects in the experimental setting; subsequent equations show how these relations are altered with changes in the values of different variables. These intermediate equations lead up to a final equation expressing the general principle. Thus the form of the proof is commonly that of a series of equations, each one following logically from the preceding one.

Proofs are typically equation-based because equations are a convenient and economical form for expressing relationships. The alternative, describing relations through written statements, is more cumbersome to process and more costly in terms of space. Furthermore, it is not always feasible to express complex relationships as written statements. However, there may be learning consequences associated with the conventional equation-based proof format. It has been shown that equations are given special status by readers studying science texts. Novices (i.e., people who are not familiar with scientific material) tend to automatically assume equations are important and thus devote extra effort to learning equations (Dee-Lucas & Larkin, 1988a,b). This tendency to focus on equations could result in learners processing individual proof equations in isolation without trying to determine why one equation follows from another, and therefore missing the logic behind the proof.

The purpose of the present research was to determine how the use of equations in proofs influences readers' understanding. This was done by comparing novices' ability to solve problems after studying equation-based proofs and proofs written out in verbal form. If equations disrupt learners' processing of proofs, then they should be less able to use proof content to solve problems when the proof uses equations instead of verbal statements. This effect would indicate that although the equation-based proof format is practical and economical, it poses some processing difficulties for learners.

For this experiment, an equation-based and verbal version of each of two passages were written. Both versions began with a paragraph stating the principle, followed by a proof of that principle. The principle was identical in each version, with the proof varying in the form. In the equation version, the proof consisted of the traditional series of related equations. In the verbal version, all equations were translated into equivalent verbal statements.

Novices read one of the two versions of each passage and then solved problems by either (1) predicting what would happen to a variable when the value of a second variable was changed, or (2) solving for a quantity. Performance was compared for the two versions to determine how the equation-based and verbal formats affected learners' understanding of the text, as indicated by their ability to use the information solve problems.

<div align="center">Method</div>

Stimulus Materials

The two passages presented Pascal's principle and the equation of continuity. They were 21 and 25 sentences in length respectively. Each passage included one diagram. The equation-based and verbal versions of one passage are presented in Appendix A.

The problems for each passage tapped three abilities: (1) direct application of the passage content, (2) transfer requiring subjects to apply the information in a new way, and (3) application of a premise or relation appearing in the proof. The following paragraphs describe these problem types more completely. Each problem-type category contained questions asking for the following kinds of responses: (1) predict how a variable will change, or (2) find the value of a quantity.

Direct application. These questions required straightforward application of the principle to a situation similar to that presented in the text. There was a direct mapping between the problem and the passage so subjects could simply match problem elements to elements presented in the text. Therefore these questions did not require subjects to have a deep understanding of either the proof or principle. For the prediction problems, subjects were to indicate the relative values of two variables, or indicate how one variable's value would change with a change in the value of a second variable (e.g., increase, decrease, or remain the same). There was one prediction problem in this category for each passage. For the value-finding problems, subjects were to solve for the new value of a variable. There was one value-finding problem for Pascal's principle and none for equation of continuity.

Transfer. These problems involved situations that differed perceptually from those in the original text (e.g., a differently shaped container of liquid) and/or dealt with a variable not explicitly addressed in the text (e.g., using Pascal's principle to find the pressure on the bottom of a tank of liquid, as opposed to a point within the liquid). Thus these problems required a more thorough understanding of the passage. Although they could be answered solely on the basis of the statement of the principle, the proof reinforced the understanding of how the principle could be applied in these transfer problems. The prediction problems involved predicting changes in variables. The value-finding problems required subjects to actually solve for a new value. There were three prediction and three value-finding problems of this type for Pascal's principle, and one prediction and value-finding problem for equation of continuity.

Proof. These problems required knowledge of premises or relations appearing in the proof (e.g., an equation for calculating a component variable in the principle, or an expression of an assumption crucial for the proof). Thus these problems tested for understanding of the proof alone. There were three prediction problems and one value-finding problem of this type for each of the two passages.

It was not feasible to develop an equal number of each problem type for each passage due to differences in the nature of the content in the two passages. There were 12 prediction problems (5 for the equation of continuity; 7 for Pascal's principle) and 7 value-finding problems (2 for the equation of continuity; 5 for Pascal's principle). For each passage, subjects were given three hypothetical situations and answered the prediction and value-finding problems based on those situations. The question set used for one passage is presented in Appendix B.

Subjects

The subjects were 40 undergraduates who had not completed more than one semester of college-level physics. This level of expertise was chosen so that subjects would be unfamiliar with the passage topics, and thus could be potentially influenced by the manner in which the content was presented.

Procedure

Subjects were told that the purpose of the experiment was to find out what types of information people find difficult and easy to understand in scientific texts. They were told that they would be reading several short science texts and solving simple problems. Subjects first read a practice passage and solved problems based on that passage. They next read the two target passages sequentially, and then solved related problems. They read one passage in the equation-based form and the other in the verbal form. The order of presentation was counterbalanced. In all cases, subjects were given as much time as they wanted to study the passages and solve the problems. They did not have access to the passages while answering the questions.

Results

The dependent measure was the number of questions of each type answered correctly. Because the dependent measure was dichotomous (correct/incorrect), the data were analyzed with a logistic regression. The variables included in the analysis were question type (direct application, transfer, or proof), response type (predict or find-value), text version (equation-based or verbal), and passage. The analysis showed that the response type interacted with all other variables, indicating different results for the prediction and value-finding problems. Therefore, separate regressions were run for these two problem sets. The results of these analyses are summarized below.

Prediction Problems

The prediction problems were analyzed in two ways. First, a regression was run on the complete data set to examine the best-fitting model for all problems of this type. Based on this initial analysis,

an additional regression was run on just the proof-prediction problems to further clarify the data pattern for this particular problem set.

Complete problem set: overall analysis. The best-fitting model for the prediction problems included the main effects of question type, version, and passage, and the version x type and type x passage interactions (χ^2=2.40, df=3, p<.493). The regression coefficients for this model are shown in Table 1.

Insert Table 1 about here

The main finding of interest is the version x type interaction, shown in Figure 1. The equation-based text produced better performance on the direct-application questions, but the verbal version resulted in better performance on the proof questions. The proportion of questions answered correctly for the direct-application questions was .72 with the verbal version and .87 with the equation version. These means for proof questions were .85 correct with the verbal version and .72 correct with the equation version. Thus the verbal version enhanced subjects' ability to answer prediction questions about the proof, and the equation version improved their ability to make predictions involving straightforward application of the principle. The equation-based version also produced a small improvement in performance on the transfer problems (proportion correct was .71 with the verbal version and .77 with the equation version).

Insert Figure 1 about here

The type x passage interaction shown in Table 1 was due to the fact that subjects found different types of questions easier in different passages. Performance was best on the proof questions with the equation of continuity passage, and on the principle questions with the passage about Pascal's principle. This interaction simply reflects content differences between passages. The proportion of questions answered correctly for the passage on Pascal's principle were .70 (direct-application), .75 (transfer), and .67 (proof). The corresponding means for equation of continuity were .90, .72, and .91.

The relatively large parameter estimate for the main effect of passage indicates that subjects performed better overall on the questions relating to the equation of continuity. The small parameter estimates for the main effect of question type indicates that there was not much difference in overall performance on the three types of questions, but subjects had somewhat more difficulty answering the transfer questions. The small estimate for the main effect of version shows that passage version did not influence overall test performance summed across question type.

Proof prediction problems: further exploration. As expected, presenting the proof in a non-equational form increased subjects' ability to use the proof content in solving prediction problems. There are two potential sources for this effect.

First, readers may have processed the quantitative relations more thoroughly when they were presented as written statements rather than equations. With the equational presentation, subjects may have tried to simply memorize the symbolic expressions without really understanding them. In contrast, the verbal presentation may have encouraged subjects to think more about the meaning of the relations, resulting in a better understanding of this content. This effect would be reflected in performance on questions pertaining to content whose presentation form varied in the two passage versions. The verbal version would produce better performance than the equational version on these questions.

Second, readers may have attended more evenly to the proof content when it was presented in verbal form and therefore performed better on proof questions because they were able to remember more of the proof. With the equational version, subjects may have focused primarily on the equations to the exclusion of non-quantitative relational content. In this case, the advantage of the verbal presentation would be to increase overall recall by increasing recall of non-quantitative content, rather than to increase depth of understanding of quantitative relations. This effect would be seen in performance on questions relating to content whose presentation form was identical in both passage versions. The verbal version would produce better performance than the equational version, indicating that the equational presentation caused readers to focus on equations at the expense of non-quantitative content.

In order better understand why the verbal presentation form improved performance on the proof questions, the question set was divided into questions assessing information that (1) varied in form across passage versions, and (2) remained constant in form across versions. Content that varied in form expressed a quantitative relation (as either an equation or its verbal equivalent). Content remaining constant in form across passage versions contained non-quantitative information (e.g., density is constant in incompressible fluids). There were three questions in each category. With correct/incorrect answer as the dependent measure, a logistic regression was run including the variables of version (equation-based/verbal) and form constancy (form constant/form varied). The results indicated significant main effects of both variables (χ^2=1.46, df=1, p<.23) and no significant interactions. The coefficients for this model are shown in Table 2.

The main effect of version with no interactions shows that for questions about both form-varied (quantitative) and form-constant (non-quantitative) content, subjects did better with the verbal than equation-based texts. This effect is shown in Figure 2. For questions about form-constant content, subjects answered .60 correct with the equation-based version and .74 correct with the verbal version. For questions about form-varied content, subjects answered .85 correct with the equation-

based version, and .97 correct with the verbal version. This suggests that the verbal presentation both improved understanding of quantitative relations (i.e., information that varied in form) and increased overall recall of proof content.

Insert Table 2 & Figure 2 about here

The main effect of targeted content (i.e., better performance on questions about form-varied content) is not interpretable because the two question categories were correlated with passage topic (i.e., all questions about form-varied content dealt with the equation of continuity, and all form-constant questions dealt with Pascal's principle). It could be that subjects found one passage easier to understand than the other, or this effect could simply reflect content differences between the two question sets independent of the passage topic.

Find-Value Problems

For questions requiring subjects to solve for a quantity, the best-fitting model included the main effects of type, version, and passage, and the version x passage interaction (χ^2=2.97, df=4, p<.563). Unlike the prediction problems, there was no version x type interaction. The parameter estimates for this model are presented in Table 3.

Insert Table 3 about here

There are two aspects of this question set that make interpretation of the data pattern problematic. First, there were relatively few find-value problems in each question category (1 direct application, 2 proof, and 4 transfer). Second, the performance level on these questions was very low (subjects averaged 2 correct out of the total 7 find-value problems, as opposed to 9 correct out of the 12 prediction problems). Because there were few questions of this type and subjects performed poorly, the assessment of the effects of the independent variables is based on very few data points.

The version x passage interaction indicates that subjects were better able to solve problems about the equation of continuity with the equation-based version, and about Pascal's principle with the verbal version. The proportion of problems solved correctly was .22 (verbal) and .35 (equation-based) for equation of continuity; for Pascal's principle these means are .37 (verbal) and .25 (equation-based). However, because there were only two questions in this problem set for the equation of continuity passage, this interaction cannot be assumed to be reliable.

The relatively large parameters for the main effect of question type shows that subjects found some types of questions easier to answer than others. Again, because of the small number of each question type (1 direct application, 2 proof, and 4 transfer), it is difficult to interpret this main effect. However, the parameter estimates suggest that subjects found the proof questions to be the most

difficult (mean proportion correct of .45 for direct application, .34 for transfer, and .15 for proof questions).

The small parameter estimate for passage version shows that overall performance on these questions was not influenced by whether the proof was equation-based or presented verbally. Similarly, the small estimate for the main effect of passage indicates that there was little difference in the difficulty level of the question sets for each passage.

<div align="center">Discussion</div>

This study examined the effects of presentation form on readers' understanding of scientific proofs. The results indicate that the comprehension of proofs is influenced by whether the content is expressed in the form of equations or written out as verbal statements.

Readers had more difficulty answering proof-related questions with the equation-based format. Previous research suggests that one reason for this may be because novice readers assume that equations are important, and therefore focus on equations to the exclusion of other types of content (Dee-Lucas & Larkin, 1986, 1988b). This would include, in the present study, the important content logically relating the equations into a coherent proof. Subjects' knowledge of non-quantitative proof information indicates that this was the case -- subjects were better able to use this content when studied in the verbal proof format. This suggests that in the equation-based proofs, readers were distracted by equations, focusing primarily on this content so that comprehension of other content suffered. This effect of equations on the comprehension of non-equational content is particularly troublesome in the case of scientific proofs, because understanding the relationships among quantitative relations is especially important for full understanding of this content. It has been shown with other content domains that assumptions and logical consequences are the types of content that readers are least likely to recall from logical arguments (Marcus, 1982). The current research suggests that students' understanding of the assumptions and overall logical structure of scientific proofs may be particularly hindered by the use of equations. This result could also explain in part the more general finding from other research that students tend to develop fragmented representations of scientific content domains (DiSessa, 1988; Labudde, Reif, & Quinn, 1988; Reif & Allen, 1989).

Because subjects focus primarily on equations in the equation-based presentation, it might be expected that they would do better on equation-related questions with this format. However, readers performed better on these questions with the verbal presentation. Previous research shows that novices recall quantitative relations better when they are presented as equations (Dee-Lucas & Larkin, 1988a). This suggests that the locus of the problem-solving effect for these questions is in applying the content rather than recalling it. That is, readers may be better able to recall information presented in equational form (as shown in previous research) but have a better understanding of that content when it is presented verbally, as indicated by the ability to use the information in problem-solving. This suggests that readers are focusing on equations for the purpose of memorizing them,

without acquiring a thorough understanding. It may be that the compact form of equations encourages students to memorize equations as a unit, and discourages students from decoding the symbolic expressions in order to fully understand them. On the other hand, presenting the information verbally allows readers to start with a semantic expression of the relationship which may be more difficult to recall, but ultimately results in a better understanding. This suggests that in understanding equation-based proofs, novice readers may tend to view the relationships among component equations in terms of how the expressions change (i.e., what symbols are added or dropped) without attending to why they change.

It was expected that an improved understanding of the proof would also result in a deeper understanding of the principle, such that performance on direct application and transfer questions would also be facilitated by the verbal version of the proof. However, this was not the case. Subjects performed better with the equation-based proof on direct-application problems, and slightly better with this presentation format on the transfer problems.

An improved understanding of the proof most likely did not contribute to performance on the direct-application questions because a deep understanding of the principle was not necessary for answering these questions successfully. Because these problems mapped directly onto the situation presented in the text, readers could simply substitute values into the relations expressed by the principle to answer these correctly. The improved performance with the equation-based presentation is probably related to differences in the form in which the principle was presented at the end of the proof with the two passage versions. In both versions, the principle was initially presented in verbal form, with symbols for quantities but with relations among quantities written out in sentence form. However, when the principle was repeated at the end of the proof, it was repeated in sentence form (similar to the initial presentation) in the verbal version and as an equation in the equation-based version. As noted earlier, there is evidence that readers recall quantitative relations better when they are presented as equations. Thus the beneficial effect of the equation-based proof on direct-application questions may be because readers were more likely to recall the principle when they had seen it as an equation at the end of the proof. Additionally, the presentation of the principle in two different forms (verbal and equation) in the equation-based version may also increase the memorability of that information over viewing it twice in sentence form.

The lack of an improvement on the transfer questions with the verbal proof form suggests that having a thorough understanding of the assumptions and logic underlying a scientific principle does not necessarily help a student understand how to apply the principle to novel situations. Subjects performed relatively well on these questions (mean of .74 correct). Although proof format did not have a strong effect on responses, performance was somewhat improved with the equation-based proofs, a response pattern matching that of the direct-application questions. The similarity in response patterns between these two question sets is probably due to the nature of the transfer

problems. These problems required readers to apply the principle to a situation that differed from that presented in the text (e.g., find pressure in an odd-shaped container of fluid). There were two bases on which these questions could be answered. First, the questions could be answered solely on the basis of the statement of the principle presented prior to the proof in the text. This was a complete statement of the principle which included all the necessary qualifications. Thus if subjects recalled this completely, they could answer the transfer questions without relying on knowledge of the proof. Second, subjects could work through the proof mentally and determine what variables were constant and what varied with size, shape, etc., and answer the question without completely recalling the qualifications presented in the statement of the principle (of course, they would still need to recall the basic relationship, e.g. area times speed is constant).

The fact that improved understanding of the proof was not related to performance on the transfer problems suggests that students do not rely on proof information in problem solving (unless the problem specifically involves the proof). The similarity in performance level and pattern of performance between the direct-application questions and the transfer questions also suggests that subjects answered the transfer questions on the basis of recall of the principle alone. As with the direct-application problems, the slightly better performance on the transfer problems with the equation-based proof may be due to enhanced recall of the principle as a result of viewing it in two different forms.

Students apparently try to solve problems using their knowledge of the principle, and do not work through the proof, even if memory for the principle fails. Additionally, better comprehension of the proof does not facilitate recall of the principle -- this is facilitated by repetition of the principle in different forms. These findings raise questions about the role of proofs in facilitating understanding of scientific principles. There are two factors that could contribute to the finding that students tend to rely on their understanding of the principle, independent of its proof, in problem solving.

First, previous research suggests that students do not always integrate proof and principle information (Dee-Lucas & Larkin, 1989). Because the principle and proof are completely comprehensible as stand-alone units, it is not necessary for readers to actively interrelate the information from the two passage segments. For example, readers might not relate the fact that Pascal's principle holds true for all points in a fluid (presented as part of the principle) to the fact that the component quantities used to calculate pressure are constant (presented as part of the proof). If this is true, then students' comprehension of the proof would not be strongly related to their ability to recall and use the principle. However, this suggests that a relationship between the comprehension of scientific proofs and principles may be found if students are given a learning set encouraging the active integration of passage content. Students may not be interrelating these two types of content because they assume that the proof is simply elaborative information which is not critical to the main point of the text (i.e., the principle). Prior research shows that novices have

strong preconceptions about what types of information are important and unimportant in science texts (Dee-Lucas & Larkin, 1988a,b). If students are encouraged to attend to and integrate proof and principle content, then a better understanding of the proof could result in a more thorough understanding of the principle, which would be reflected in better problem-solving ability.

Second, it is also possible that the transfer problems included in this study were not difficult enough to prompt subjects to draw on their knowledge of the proof. Subjects may be more likely to rely on this knowledge in solving more complicated problems, in which a variety of elements differ between the original text and the problem situation. For these types of problems, a better understanding of the rationale underlying the principle (i.e., the proof) could help in determining whether the principle applies, and how it would apply to the problem setting. In general, the effectiveness of elaborative information in texts (i.e., information that supports or clarifies the main points) has been found to be highly specific to the nature of the elaborations and the skill being learned. For example, Reder, Charney, and Morgan (1986) found that in teaching how to use a personal computer, text elaborations illustrating the syntax of operating commands facilitated learning, whereas elaborations explaining basic concepts and their applicability did not. Similarly, Pirolli and Anderson (1985) found that for facilitating the learning of a computer language, providing examples illustrating how to perform a task is better than examples clarifying the outcome of a procedure. It may be that for facilitating simple scientific problem-solving, providing examples of problems may be more beneficial than presenting proofs (Ross, 1987). Proofs, on the other hand, could prove beneficial in the case of more complex problem situations.

Although a better understanding of the proof did not aid in solving transfer problems of the type provided in this study, there are other reasons for teaching students proofs. Proofs serve important learning functions besides improving problem solving. Scientific proofs illustrate scientific methodology and in some instances provide information as to the historical context and significance of the discovery of a scientific relation. Furthermore, an understanding of the nature and types of assumptions underlying a scientific principle can be important in and of itself, both in showing how one constructs a logical argument and in promoting understanding of other scientific relations sharing assumptions.

All of the results in the current study were based on the prediction problems. The performance on the problems requiring readers to solve for a quantity was too poor to yield enough data for reliable results. It may be that subjects did not spend enough time with the material in this study to be able to solve this type of question -- they were able to predict how a change in one variable would affect another, but could not remember the details needed to actually calculate the value change. However, the low performance on these questions is consistent with general findings from research on scientific problem-solving indicating that students lack the skills necessary for solving quantitative problems (Heller & Reif, 1984; Larkin, McDermott, Simon,& Simon, 1980).

Summary

These findings show that the manner in which a proof is presented influences students' ability to understand and recall it. The equation-based format hindered comprehension both by reducing the comprehensibility of quantitative relations and by distracting students from non-quantitative information. These results suggest that the traditional format for presenting proofs is not optimal for enhancing learning. This is in part because novice readers use presentation form as a guide for assessing importance as they are reading. Novices assume that equations signal important content, whereas equations are typically used as a convenient notation for expressing any quantitative relation. Additionally, equations appear to be more difficult for novices to understand than their verbal equivalents. Equations require complex processing in order to decode the symbolic expressions into their semantic representation. They may be easier to recall, but the results of this research suggest that they are harder to understand.

In general, this research indicates that novice students need guidance in processing scientific proofs to aid them in understanding and interrelating the content. This guidance can be provided in part by limiting the use of equation notation to critical proof content (helping to call attention to this content), and by elaborating on equations to aid novices with the decomposition into a semantic representation. Additionally, important non-quantitative information should be presented in a manner that calls attention to it, either by signally its importance through typographic cues (e.g., italicizing, underlining, etc.), or by explicitly indicating its importance in the text (e.g., "you should note that," "it is important to understand," etc.).

This research also raises questions about the role of proofs in facilitating understanding of scientific principles. There was no evidence in this study that a deeper understanding of the proof contributed to a deeper understanding of the principle. This may be because of the nature of the problems included in this research -- more difficult problems may in fact demonstrate such a relationship. However, it is also possible that novice readers fail to process proofs in a manner that interrelates them with their corresponding principle. If so, then novices may need additional help in determining the relationship between the quantitative relations expressed by the principle and the related properties of the component quantities presented in the proof.

References

Dee-Lucas, D. & Larkin, J.H. (1986). Novice strategies for processing scientific texts. *Discourse Processes*, **9**, 329-354.

Dee-Lucas, D. & Larkin, J.H. (1988a). Attentional strategies for studying scientific texts. *Memory & Cognition*, **16**, 469-479.

Dee-Lucas, D. & Larkin, J.H. (1988b). Novice rules for assessing importance in scientific texts. *Journal of Memory & Language*, **27**, 288-308.

Dee-Lucas, D. & Larkin, J.H. (1989). *Organization and comprehensibility in scientific proofs, or "Consider a particle p..."* (Tech Report ACP #19). Carnegie Mellon University.

DiSessa, A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, N.J.: Erlbaum.

Heller, J.I. & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition & Instruction*, **1**, 177-216.

Labudde, P., Reif, F., & Quinn, L. (1988). Facilitation of scientific concept learning by interpretation procedures and diagnosis. *International Journal of Science Education*, **1**, 81-98.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Models of competence in solving physics problems. *Cognitive Science*, **4**, 317-345.

Marcus, S. L. (1982). Recall of logical argument lines. *Journal of Verbal Learning and Verbal Behavior*, **21**, 549-562.

Pirolli, P. L. & Anderson, J.R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology*, **39**, 240-272.

Reder, L. M., Charney, D.H., & Morgan, K.I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition*, **14**, 64-78.

Reif, F. & Allen, S. (1989). *Interpreting and teaching scientific concepts: The example of "acceleration."* (Tech Report, EMST, School of Education). University of California, Berkeley.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 629-639.

Table 1

Coefficients for the Logistic Regression Model for the Prediction Problems

Variable	Coeff.	St. Error	Coeff./St. Error
Type:			
Proof	.148	.183	.808
Direct Application	.207	.242	.859
Transfer	-.355	.188	-1.885
Version:	-.089	.132	-.671
Passage:	.486	.142	3.421
Version x Type:			
Proof	.509	.165	3.077
Direct Application	-.433	.222	-1.953
Transfer	-.076	.169	-.449
Type x Passage:			
Proof	.335	.179	1.876
Direct Application	.215	.234	.921
Transfer	-.550	.185	-2.969

Table 2

Coefficients for the Logistic Regression Model for the Proof Prediction Problems

Variable	Coeff.	St. Error	Coeff./St. Error
Version:	.420	.172	2.445
Form Constancy:	.821	.188	4.364

Table 3

Coefficients for the Logistic Regression Model for the Value-Finding Problems

Variable	Coeff.	St. Error	Coeff./St. Error
Type:			
Proof	-.983	.260	-3.776
Direct Application	.765	.262	2.923
Transfer	.219	.188	1.165
Version:	-.016	.153	-.107
Passage:	.173	.164	1.054
Version x Passage:	-.315	.153	-2.059

Figure Captions

Figure 1. Effect of passage version (verbal and equation) for three problem types (proof, direct application, and transfer).

Figure 2. Effect of passage version (verbal and equation) when the form of the content assessed by the question was constant (qualitative information) and varied (quantitative information) across passage version.
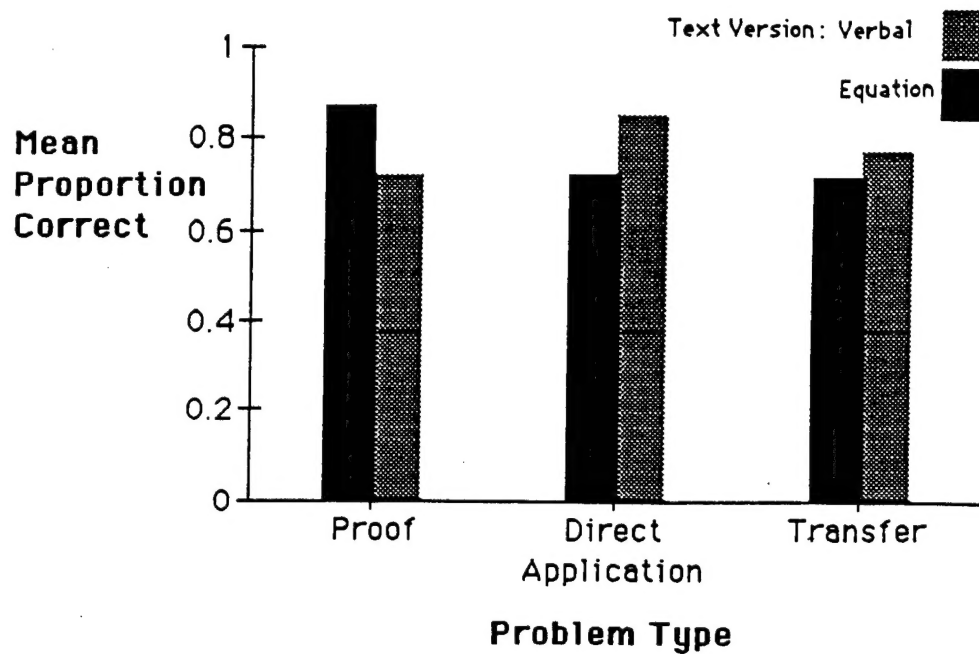
Figure 1. Effect of passage version (verbal and equation) for three problem types (proof, direct application, and transfer).
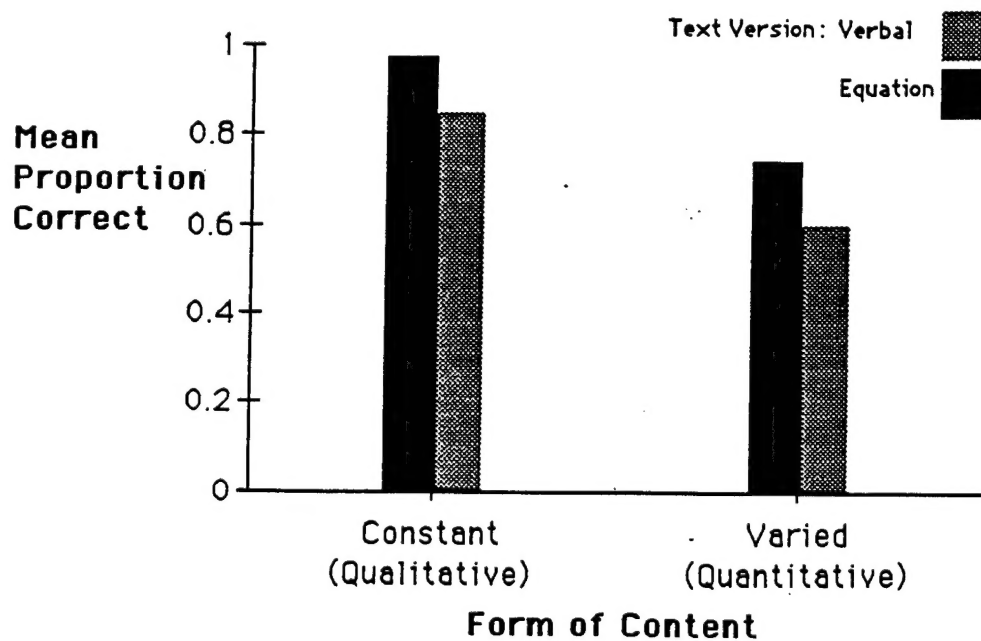

Figure 2. Effect of passage version (verbal and equation) when the form of the content assessed by the question was constant (qualitative information) and varied (quantitative information) across passage version.

Appendix I
Sample Passage showing Alternate Wording for Verbal and Equation Versions

Pascal's Principle - Pressure in a Fluid

The term "fluid statics" or "hydrostatics" is applied to the study of fluids at rest. In fluid statics, we assume that the fluid and any other relevant objects, such as containers, are in equilibrium. The laws of hydrostatics describe the forces produced by fluids and how they interact when fluids are at rest.

Pascal's principle describing how pressure varies in a fluid with changes in external pressure was discovered in the seventeenth century by Blaise Pascal. Pascal's principle states that the added pressure applied anywhere on a fluid, Δp, is equal to the corresponding change in pressure at any point in the fluid, $p_2 - p_1$. This is a necessary consequence of the laws of fluid mechanics, as shown below.

Assume that there is a liquid in equilibrium in a closed cylinder that is fitted with a piston. [Figures placed approximately here in experimental text.]

Continuation of Verbal Version	*Continuation of Equation Version*
Using the piston, we apply an external pressure to the top of the fluid and measure the pressure at an arbitrary point in the liquid (see figure 1).	Using the piston, we apply an external pressure p_e to the top of the fluid and measure the pressure p_1 at an arbitrary point A in the liquid (see figure 1).
We now increase the external pressure by an arbitrary amount (see figure 2).	We now increase the external pressure p_e by an arbitrary amount Δp (see figure 2).
We can calculate the size of the pressure change by subtracting the old pressure at that point from the new pressure at that point.	We can calculate the size of the pressure change, Δp, by subtracting the old pressure at that point, p_1, from the new pressure at that point, p_2.
The old pressure is equal to the sum of the original external pressure and the pressure due to the liquid above the point.	The old pressure p_1 is equal to $$p_1 = p_e + p_l$$ where p_e is the original external pressure and p_l is the pressure due to the liquid above the point.
The new pressure is equal to the sum of the original external pressure, the pressure due to the liquid above the point, and the increase in the external pressure (over the original external pressure).	The new pressure p_2 is equal to $$p_2 = \Delta p + p_e + p_l$$ where Δp is the increase in the external pressure (over the original external pressure p_e).
When the external pressure is increased, the original external pressure and the pressure due to the liquid above the point do not change.	When the external pressure p_e is increased, the original external pressure p_e and the pressure p_l due to the liquid above point A do not change.
The fluid pressure above the point is equal to the product of the fluid density, the distance of the point from the top of the fluid, and the gravitational constant.	The fluid pressure above the point is equal to $$p_l = \rho g h$$ where ρ is the fluid density, h is the distance of the point from the top of the fluid, and g is the gravitational constant.

This quantity remains the same when pressure is added because liquids are incompressible—thus the distance of the point and the fluid density are not altered by the added pressure, nor is the gravitational constant.

This quantity p_l remains the same when pressure is added because liquids are incompressible—thus h and ρ are not altered by the added pressure, nor is the gravitational constant g.

Because the original external pressure and fluid pressure above the point remain unchanged when pressure is added, subtracting the equation for old pressure from the equation for new pressure leaves only the added external pressure.

Because the original external pressure p_e and fluid pressure p_l above the point remain unchanged when pressure is added,
$$(\Delta p + p_e + p_l) - (p_e + p_l) = \Delta p .$$

Thus the change in pressure at an arbitrary point in a fluid is equal to the change in external pressure.

Thus,
$$p_2 - p_1 = \Delta p .$$

This is the result known as Pascal's Principle.

Although we assume that liquids are incompressible, they are in fact slightly compressible. This means that a change of pressure applied to one portion of a liquid propagates through the liquid as a wave at the speed of sound. Once this disturbance has died out and equilibrium is established, it is found that Pascal's Principle is valid.

Appendix II
Test for Pascal's Principle

This test included three diagrams, with several questions about each. For questions about the first two diagrams (questions sets I and II below), each question was followed by the options:
    a. insufficient information given
    b. no
    c. yes (if sufficient information is given, indicate the size of the change _____)

These questions were classified according to the nature of the response:
    (1) predict whether a change occurs (i.e., select from options a, b, and c)
    (2) find the value of a change (i.e, calculate the value if selecting option c)

The last question (question set III below) required subjects only to solve for a value .

The questions were also categorized according to the kind of ability they tapped:
    direct application of the passage content [dir app]
    transfer requiring subjects to apply the information in a new way [trans]
    application of some component of the proof [proof]
These categories are indicated next to the questions below, but this information was not provided to the subjects taking the test.

Problems for Pascal's Principle
I. Assume we have a 10 gallon tank filled with an incompressible oil at equilibrium. A random point within the tank is labelled A. The top of the tank is fitted with a piston. Using the piston, we increase the pressure on the oil by 5 pounds.

    Does the pressure at point A change? (circle one) [dir app]
    Does the pressure on the bottom of the oil tank change? (circle one) [trans]
    Does the density of the oil change? (circle one) [proof]
    Does the pressure due to the oil above the point A change? (circle one) [proof]
    Does the distance to the point A from the top of the oil change? (circle one) [proof]

II. Consider the apparatus shown above. We have a container fitted with a piston. There is a long tube extending from the side of the container. The container and the tube are filled with mercury. Two points within the mercury are labelled A and B. The top of the tube is fitted with a cap and labelled C. The mercury is incompressible and at equilibrium. Suppose we use the piston to increase the pressure on the mercury by 1.7 pounds.

    Does the pressure at the point B change? (circle one) [trans]
    Does the pressure on the cap of the tube changed? (circle one) [trans]

III. We have a test tube filled with incompressible sulfuric acid. The sulfuric acid is in equilibrium. We know the following information:
    The atmospheric pressure on top of the test tube is 40,000.
    The cross-sectional area of the test tube is 1.
    The density of the sulfuric acid is 1.05.
    The gravitational constant is 980.
    The height from the top of the sulfuric acid to the point A is 10.
    The pressure on the bottom of the test tube is 307,000.
    The volume of sulfuric acid in the test tube is 100.

    What is the pressure at the labelled point A in the test tube? [proof]